

# **Wirkung und Wirkmechanismen von regelmäßigem Schülerfeedback in der Sekundarstufe**

**Eine explorative Untersuchung im Physikunterricht**

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Simon Rösch

aus Weil am Rhein, Deutschland

Weil am Rhein, 2017

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel  
[edoc.unibas.ch](http://edoc.unibas.ch)

---

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Nikolaus J. Kuhn

Prof. Dr. Peter Labudde

Prof. Dr. Kai Niebert

Basel, den 21. März 2017

Prof. Dr. Martin Spiess (Dekan)

## **INHALT**

<b>1</b>	<b><u>EINLEITUNG</u></b>	<b>9</b>
<b>2</b>	<b><u>FORSCHUNGSSTAND</u></b>	<b>11</b>
<b>2.1</b>	<b>VOM FEEDBACK ZUM SCHÜLER-FEEDBACK: CHARAKTERISTIK UND DEFINITIONEN</b>	<b>11</b>
2.1.1	PERSONALE EINFLUSSFAKTOREN AUF DIE FEEDBACKVERARBEITUNG	14
2.1.2	RELIABILITÄT UND VALIDITÄT VON SCHÜLERBEURTEILUNGEN	18
2.1.3	WEITERE MÖGLICHE VORTEILE VON SCHÜLERFEEDBACK	24
2.1.4	SELBSTWAHRNEHMUNG DES EIGENEN UNTERRICHTS	25
<b>2.2</b>	<b>WIRKUNGEN VON SCHÜLERFEEDBACK</b>	<b>26</b>
2.2.1	AUSGEWÄHLTE UNTERSUCHUNGEN AN HOCHSCHULEN	26
2.2.2	AUSGEWÄHLTE UNTERSUCHUNGEN IM REGELSCHULBEREICH	28
<b>3</b>	<b><u>ENTWICKLUNG EINES SCHÜLER-KURZFEEDBACKS ALS MESS- UND INTERVENTIONSINSTRUMENT</u></b>	<b>33</b>
<b>3.1</b>	<b>LERNEN ALS KONSTRUKTIONSPROZESS</b>	<b>33</b>
3.1.1	VOM RADIKAL-KONSTRUKTIVISTISCHEN ZUM SOZIAL-KONSTRUKTIVISTISCHEN LERNVERSTÄNDNIS	33
3.1.2	DIE ROLLE VON EXPERIMENTEN BEIM KONSTRUIEREN NEUEN WISSENS	35
<b>3.2</b>	<b>DER KURZFEEDBACKBOGEN</b>	<b>37</b>
<b>4</b>	<b><u>DIE STUDIE</u></b>	<b>41</b>
<b>4.1</b>	<b>FORSCHUNGSFRAGEN</b>	<b>41</b>
<b>4.2</b>	<b>INSTRUMENTE ZUR ERFASSUNG BEEINFLUSSBARER FAKTOREN UND IHRE THEORETISCHE FUNDIERUNG</b>	<b>42</b>
4.2.1	FAKTOREN DER UNTERRICHTSQUALITÄT	44
4.2.2	UNTERRICHTSENTWICKLUNG/UNTERRICHTSHANDELN DER LEHRPERSON	46
4.2.3	LEHRPERSONEN-EINSCHÄTZUNG DER SOZIAL-KOGNITIVEN VORAUSSETZUNGEN UND DES BEURTEILUNGSVERMÖGENS VON UNTERRICHT DURCH SCHÜLERINNEN UND SCHÜLERN	49
<b>4.3</b>	<b>INSTRUMENTE ZUR ERFASSUNG DER EINSCHÄTZUNG DES SCHÜLER-KURZFEEDBACK-VERFAHRENS</b>	<b>50</b>
<b>4.4</b>	<b>FORSCHUNGSDESIGN</b>	<b>51</b>
4.4.1	ABLAUF DER HAUPT-STUDIE	53
<b>4.5</b>	<b>AUSWERTUNGSMETHODIK</b>	<b>54</b>

---

4.5.1	QUALITATIVE INHALTSANALYSE DER INTERVIEWS	54
4.5.2	TRIANGULATION DER DATEN	57
<b>5</b>	<b>ERGEBNISSE</b>	<b>59</b>

---

<b>5.1</b>	<b>STATISTISCHE GÜTE DER SKALEN</b>	<b>59</b>
5.1.1	RELIABILITÄT DER SKALEN ZUR UNTERRICHTSQUALITÄT AUS SCHÜLERSICHT	59
5.1.2	RELIABILITÄT DER SKALEN DES SCHÜLER-KURZFEEDBACKS	60
<b>5.2</b>	<b>ERGEBNISSE BEZOGEN AUF DIE GESAMTE STICHPROBE</b>	<b>62</b>
5.2.1	DEMOGRAPHISCHE DATEN DER SCHÜLERINNEN UND SCHÜLER	62
5.2.2	DEMOGRAPHISCHE DATEN DER LEHRPERSONEN	64
5.2.3	UNTERRICHTSQUALITÄT AUS SCHÜLERSICHT	65
5.2.4	WAHRNEHMUNG DES SCHÜLER-KURZFEEDBACK-VERFAHRENS AUS SCHÜLERSICHT	67
5.2.5	VERLAUF UND TENDENZ DER SCHÜLERKURZFEEDBACKS ÜBER ALLE SCHULKLASSEN	73
5.2.6	DIE SCHRIFTLICHE BEFRAGUNG DER LEHRPERSONEN	75
5.2.7	DIE INTERVIEWS MIT DEN LEHRPERSONEN	78
5.2.8	ZUSAMMENFASSUNG DER ERGEBNISSE ÜBER DIE GESAMTE STICHPROBE	88
<b>5.3</b>	<b>INTERKLASSENVERGLEICH UND GRUPPENBILDUNG</b>	<b>90</b>
5.3.1	INTERKLASSENVERGLEICH, SKALEN DER UNTERRICHTSQUALITÄT	90
<b>5.4</b>	<b>INTERGRUPPENVERGLEICH</b>	<b>95</b>
5.4.1	INTERGRUPPENVERGLEICH, VERLAUF SCHÜLER-KURZFEEDBACKS	95
5.4.2	INTERGRUPPENVERGLEICH, EINSCHÄTZUNG DES SCHÜLER-KURZFEEDBACK-VERFAHRENS DURCH DIE SCHÜLERINNEN UND SCHÜLER	105
5.4.3	INTERGRUPPENVERGLEICH, FRAGEBOGENERGEBNISSE DER LEHRPERSONEN	107
5.4.4	ZUSAMMENFASSUNG DER INTERGRUPPENVERGLEICHE	108
<b>5.5</b>	<b>ANALYSE DER EINZELNEN GRUPPEN</b>	<b>109</b>
5.5.1	GRUPPE 1 (KLASSE 10)	110
5.5.2	GRUPPE 2 (KLASSEN 7 UND 8)	113
5.5.3	GRUPPE 3 (KLASSEN 2, 4 UND 6)	116
5.5.4	GRUPPE 4 (KLASSEN 1, 5 UND 9)	118
5.5.5	GRUPPE 5 (KLASSE 3)	121
5.5.6	ZUSAMMENFASSUNG DER EINZELANALYSEN	125
<b>6</b>	<b>DISKUSSION, HYPOTHESENBILDUNG UND AUSBLICK</b>	<b>126</b>

---



<b>6.1</b>	<b>INSTRUMENTE UND DESIGN</b>	<b>126</b>
<b>6.2</b>	<b>LASSEN SICH HINWEISE IDENTIFIZIEREN BEZÜGLICH DES EINFLUSSES VON SCHÜLER-KURZFEEDBACK AUF A) AUSGEWÄHLTE SKALEN DER UNTERRICHTSQUALITÄT UNTER BESONDERER BERÜCKSICHTIGUNG DES SCHÜLER-LEHRPERSONEN-VERHÄLTNISSES UND B) DIE UNTERRICHTSENTWICKLUNG UND DAS UNTERRICHTSHANDELN DER LEHRPERSON? (FF 1)</b>	<b>128</b>
<b>6.3</b>	<b>WELCHE ENTWICKLUNGEN NEHMEN SCHÜLERINNEN UND SCHÜLER IM UNTERRICHT INNERHALB DES SCHÜLER-KURZFEEDBACK-VERFAHRENS WAHR? (FF2)</b>	<b>132</b>
<b>6.4</b>	<b>LASSEN SICH MUSTER IDENTIFIZIEREN, DIE EINE POSITIVE WIRKUNG VON SCHÜLERFEEDBACK AUF DIE UNTERRICHTSQUALITÄT BEGÜNSTIGEN KÖNNTEN? (FF 3)</b>	<b>134</b>
<b>6.5</b>	<b>WIE SCHÄTZEN LEHRPERSONEN UND SCHÜLERINNEN UND SCHÜLER DIE PRAKTIKABILITÄT DES SCHÜLER-KURZFEEDBACK-VERFAHRENS EIN? (FF 4)</b>	<b>137</b>
<b>6.6</b>	<b>WEITERE RELEVANTE ASPEKTE DES SCHÜLER-KURZFEEDBACK-VERFAHRENS</b>	<b>140</b>
<b>6.7</b>	<b>RÜCKBLICK UND AUSBLICK</b>	<b>141</b>
6.7.1	FOLGERUNGEN FÜR DIE NATURWISSENSCHAFTSDIDAKTIK	141
6.7.2	FOLGERUNGEN FÜR DIE SCHULPRAXIS	142
<b>7</b>	<b>LITERATURVERZEICHNIS</b>	<b>144</b>
<b>8</b>	<b>ANHANG</b>	<b>155</b>
<b>8.1</b>	<b>ERHEBUNGSINSTRUMENTE</b>	<b>155</b>
8.1.1	DER SCHÜLER-KURZFEEDBACK-BOGEN	155
8.1.2	SCHÜLER-FRAGEBOGEN PRÄERHEBUNG	157
8.1.3	SCHÜLER-FRAGEBOGEN; ZUSÄTZLICHE ITEMS DER POSTERHEBUNG	162
8.1.4	LEHRPERSONEN-FRAGEBOGEN, ITEMS DER PRÄERHEBUNG	163
8.1.5	LEHRPERSONEN-FRAGEBOGEN; ZUSÄTZLICHE ITEMS DER POSTERHEBUNG	166
<b>8.2</b>	<b>EINVERSTÄNDNISERKLÄRUNG DER ELTERN</b>	<b>167</b>
<b>8.3</b>	<b>ERGEBNISAUSWERTUNG DER SCHÜLER-KURZFEEDBACKS FÜR DIE LEHRPERSONEN</b>	<b>168</b>
<b>8.4</b>	<b>AUSWERTUNG DER INTERVIEWS</b>	<b>170</b>

---

## **ABBILDUNGSVERZEICHNIS**

<i>Abbildung 1: Prä-/Postvergleich Skalen der Unterrichtsqualität, Untersuchung Ditton und Arnold (2004).</i>	29
<i>Abbildung 2: Forschungsdesign</i>	43
<i>Abbildung 3: Ablaufmodell strukturierender Inhaltsanalyse.</i>	55
<i>Abbildung 4: Mittelwertvergleich Unterrichtsqualität aus Schülersicht über alle Klassen aggregiert.</i>	67
<i>Abbildung 5: Messzeitpunkte 1-4, Verlauf Kognitive Passung (keine Überforderung) über alle Schulklassen mit linearer Regression.</i>	73
<i>Abbildung 6: Messzeitpunkte 1-4, Verlauf Kognitive Passung (nicht unterfordert) über alle Schulklassen mit linearer Regression.8</i>	73
<i>Abbildung 7: Messzeitpunkte 1-4, Verlauf Kognitiv aktivierende und strukturierende Elemente über alle Schulklassen mit linearer Regression.</i>	74
<i>Abbildung 8: Messzeitpunkte 1-4, Verlauf Disziplin über alle Schulklassen mit linearer Regression (rot gestrichelt).</i>	74
<i>Abbildung 9: Messzeitpunkte 1-4, Verlauf Experimentieren als theoriegeleitetes Handeln über alle Schulklassen mit linearer Regression.</i>	74
<i>Abbildung 10: Dendrogramm, Clusteranalyse Effektstärken und Präerhebung der Unterrichtsqualität.</i>	94
<i>Abbildung 11: Offene Antworten im Schüler-Kurzfeedback: Wiederholung (sachlich)</i>	100
<i>Abbildung 12: Offene Antworten im Schüler-Kurzfeedback: Allgemein (positiv)</i>	101
<i>Abbildung 13: Offene Antworten im Schüler-Kurzfeedback: Allgemein (negativ)</i>	101
<i>Abbildung 14: Offene Antworten im Schüler-Kurzfeedback: Wiederholung (sachfern)</i>	102

## **TABELLENVERZEICHNIS**

<i>Tabelle 1: Klassifikationsschema von Attributionen nach Weiner (1985)</i>	14
<i>Tabelle 2: Auswertungskategoriensystem der Interviews</i>	56
<i>Tabelle 3: Reliabilitätsstatistik der ausgewählten und angepassten Skalen zur Unterrichtsqualität aus Schülersicht</i>	60
<i>Tabelle 4: Reliabilitätsstatistik Skalen der Schülerkurzfeedbacks</i>	60
<i>Tabelle 5: Faktorenanalyse zur Dimensionsreduktion in allgemeindidaktisches Handeln</i>	61
<i>Tabelle 6: Faktorenanalyse zur Dimensionsreduktion in Experimentieren als theoriegeleitetes Handeln</i>	61
<i>Tabelle 7: Reliabilitätsstatistik der aggregierten Skalen des Schülerkurzfeedbacks</i>	62
<i>Tabelle 8: Demographische Daten der Schülerinnen und Schüler</i>	63
<i>Tabelle 9: Demographische Daten der Lehrpersonen.</i>	64
<i>Tabelle 10: Demographische Daten der Lehrpersonen (Fortsetzung).</i>	65
<i>Tabelle 11: Mittelwertvergleich Unterrichtsqualität aus Schülersicht über alle Klassen aggregiert.</i>	66

---

<i>Tabelle 12: Veränderung des Items V2SWSF02 (alt zu neu)</i>	68
<i>Tabelle 13: Deskriptive Analyse der Items V2SWSF01, V2SWSF03, V2SWSF04, V2SWSF05, V2SFEE01 und V2SFEE02</i>	68
<i>Tabelle 14: Item V2OFRA05 Anzahl der Nennungen</i>	69
<i>Tabelle 15: Item V2OFRA06 Anzahl der Nennungen</i>	70
<i>Tabelle 16: Deskriptive Analyse Item V2SWSF02 (neu)</i>	70
<i>Tabelle 17: Item SWSF21 Anzahl der Nennungen</i>	71
<i>Tabelle 18: Deskriptive Analyse Item V2SWSF22</i>	71
<i>Tabelle 19: Korrelation zwischen wahrgenommener Veränderung des Unterrichts und Differenz der Unterrichtsqualität (Prä-Post)</i>	72
<i>Tabelle 20: Einschätzung der Lernvoraussetzungen der Schülerinnen und Schüler, des eigenen Klassenmanagements, der Homogenität und des Schüler-Lehrpersonen-Verhältnisses durch die Lehrpersonen.</i>	75
<i>Tabelle 21: Selbsteinschätzung der eingesetzten didaktischen Unterrichtsstrategien durch Lehrperson, Anzahl der Nennungen über die gesamte Stichprobe</i>	77
<i>Tabelle 22: Lehrereinschätzung: Beurteilungsvermögens von Schülerinnen und Schüler bezüglich Unterricht</i>	78
<i>Tabelle 23: Interview; Verhalten im Rahmenmodell der Unterrichtsentwicklung (allgemein und speziell); Anzahl der Nennungen über die gesamte Stichprobe</i>	79
<i>Tabelle 24: Interview; Entwicklung, wie Lehrpersonen die Schülerinnen und Schüler wahrnehmen</i>	80
<i>Tabelle 25: Einschätzung des Schüler-Feedbackverfahrens</i>	86
<i>Tabelle 26: Interklassenvergleich Skalen der Unterrichtsqualität, sortiert nach der durchschnittlichen Effektstärke d</i>	91
<i>Tabelle 27: Unterrichtsqualität, Vergleich der durchschnittlichen Effektstärke mit den Werten zur Präerhebung</i>	93
<i>Tabelle 28: Clusteranalyse</i>	94
<i>Tabelle 29: Verläufe der Kurzfeedbacks nach Gruppen getrennt.</i>	96
<i>Tabelle 30: Clustervergleich Anzahl der Nennungen auf die offene Frage des Items V2SWSF21</i>	105
<i>Tabelle 31: Allgemeine Beurteilung des Verfahrens, Item V2OFRA06</i>	106
<i>Tabelle 32: Allgemeine Beurteilung des Verfahrens, Item V2OFRA05</i>	107
<i>Tabelle 33: Einschätzung der kognitiven und sozialen Voraussetzungen, des Schüler-Lehrpersonen-Verhältnisses und des Beurteilungsvermögens der Schülerinnen und Schüler von Unterricht, Präerhebung</i>	108
<i>Tabelle 34: Antworthäufigkeiten der Kategorien in den Interviews pro Lehrperson innerhalb der Gruppe</i>	110
<i>Tabelle 35: Varianzen der Schüler-Kurzfeedbacks in den Gruppen 1-5</i>	123
<i>Tabelle 36: Varianzen der Prä-/Posterhebung Unterrichtsqualität der Gruppen 1-5</i>	123

---

## Danksagung

Das Verfassen einer Dissertation ist ohne die Unterstützung von zahlreichen hilfsbereiten Menschen nicht möglich. Ihnen möchte ich in den folgenden Zeilen danken.

Zunächst danke ich den Mitarbeitern des Zentrums für Naturwissenschafts- und Technikdidaktik. Zuallererst ist hier mein Doktorvater Prof. Dr. Peter Labudde zu nennen, der mir für die Bearbeitung des Dissertationsthema reichlich Vertrauen entgegengebracht und mich fachlich unterstützt hat. Ein sehr großer Dank geht an meinen Betreuer Dr. Matthias von Arx, der mich durch die analytischen, motivierenden und freundlichen Rückmeldungen in jeder Phase des Forschungsvorhabens dazu anregte, die Arbeit zu optimieren. Ebenfalls einen großen Dankeschön geht an Katja Stähli, die mir während der Studie half, die Daten auszuwerten und die mitunter unangenehme Arbeit der Interviewtranskription sehr zuverlässig und diszipliniert durchführte. Außerdem möchte ich Heli Schaffter danken, die mich während der Interviewauswertung tatkräftig unterstützte. Des weiteren geht mein Dank an Dr. Monika Hohlmeier und Dr. Alexander Koch, die mir in statistischen Fragen stets hilfsbereit zur Seite standen. Allen anderen Mitarbeitern, die stets ein offenes Ohr hatten und mich bei vielen kleinen oder größeren Problemen und Fragestellungen berieten, gilt ebenfalls mein Dank.

Da eine solch große Arbeit auch durch das private Umfeld beeinflusst wird und dieses wiederum beeinflusst, danke ich auch meinen Eltern, die mich in allem was ich bisher im Leben gemacht habe, unterstützt haben. Für das Korrekturlesen bedanke ich mich außerdem bei Eva-Maria Zuberer, Dirk Nagel und meiner Freundin Carolin Eichin, die mich des Weiteren vor allem während des Schreibprozesses sehr stark unterstützt und auch ertragen hat.

### Natur und Kunst

*Natur und Kunst, sie scheinen sich zu fliehen,  
Und haben sich, eh' man es denkt, gefunden;  
Der Widerwille ist auch mir verschwunden,  
Und beide scheinen gleich mich anzuziehen.*

*Es gilt wohl nur ein redliches Bemühen!  
Und wenn wir erst in abgemessnen Stunden;  
Mit Geist und Fleiß uns an die Kunst gebunden,  
Mag frei Natur im Herzen wieder glühen.*

*So ist's mit aller Bildung auch beschaffen:  
Vergebens werden ungebundne Geister  
Nach der Vollendung reiner Höhe streben.*

*Wer Großes will, muss sich zusammenraffen:  
In der Beschränkung zeigt sich erst der Meister,  
Und das Gesetz nur kann uns Freiheit geben.*

Johann Wolfgang von Goethe

## 1 Einleitung

„Man kann einen Menschen nichts lehren, man kann ihm nur helfen es in sich selbst zu entdecken.“ Diese rund 400 Jahre alte Aussage von Galileo Galilei bringt die Kunst und gleichzeitig die Problematik des Lehrens auf den Punkt. Eine Lehrperson kann die Schülerinnen und Schüler nur zum Denken anregen und sie während des Denkens auf die richtige Spur bringen, den Akt des Denkens selbst und das Beschreiten des richtigen Pfades müssen jedoch allein die Schülerinnen und Schüler vollbringen. Inwieweit dies allerdings während des Unterrichts vollzogen wird, bleibt der Lehrperson oft verborgen, denn Denk- und Lernprozesse sind während des Unterrichts schwer sichtbar zu machen. Doch genau das „Sichtbarmachen“ ist eine zentrale Forderung des Bildungsforschers John Hattie (2007). Hatties viel zitiertes Werkzeug hierfür ist das Feedback. Darunter versteht er nicht nur die reine Rückmeldung, die eine Lehrperson einem Lernenden auf eine erbrachte Leistung oder Handlung gibt, sondern vor allem das kontinuierliche Erfragen des Lernstandes bei den Schülerinnen und Schülern, das Rückmelden des Lernziels und das Anregen der Schülerinnen und Schüler, zu formulieren, was sie zu tun haben, um dem Lernziel näher zu kommen. Feedback ist also eine wechselseitige Kommunikation über den Lernprozess und nicht eine einseitige Rückmeldung von Lehrpersonenseite. Dies beinhaltet auch, dass die Lehrpersonen ihren eigenen Unterricht ständig hinterfragen, beziehungsweise hierauf bezogene Informationen erhalten sollten, um diesen an die Bedürfnisse der Schülerinnen und Schüler anzupassen.

Dieser Anpassungsprozess kann zahlreiche naturwissenschaftsdidaktische Facetten abdecken. Angefangen bei der Vorbereitung des Unterrichts, und der Frage, ob die Stoffauswahl und dessen didaktische Rekonstruktion an die sozialen, kognitiven und affektiven Voraussetzungen der Schülerinnen und Schüler angepasst sind. Bis hin zu einer gut strukturierten und interesselockenden Unterrichtsinszenierung, welche bei den Schülerinnen und Schülern einen zielorientierten Lernprozess anregen soll. Dies kann sich beispielsweise in der Auswahl der eingesetzten Experimente und deren sinnvolle Einbettung in den Lernprozess und in einem hohen Alltagsbezug bemerkbar machen. Insbesondere vor dem Hintergrund, dass der Physikunterricht zu den unbeliebtesten Schulfächern gehört (vgl. Merzyn 2008, Börlin, Beerenwinkel & Labudde 2014), scheint eine Information über den gehaltenen Unterricht wichtig zu sein, um zu erfahren, welche Unterrichtsstrategien sich bei den Schülerinnen und Schülern als verbesserungswürdig erwiesen haben, um daraufhin qualitative Veränderungen im Unterrichtsverlauf zu initiieren.

An diesem Punkt setzt das Forschungsvorhaben dieser Doktorarbeit an. Zunächst soll ein Verfahren entwickelt werden, welches ermöglicht, auf einfache Weise und in regelmäßigen Abständen Informationen über den eigenen Unterricht zu erhalten. Des Weiteren soll untersucht werden, inwieweit sich dieses Schülerfeedback aus subjektiver Lehrpersonensicht auf die Un-

terrichtsentwicklung auswirkt, und ob empirische Hinweise auf eine Veränderung der Unterrichtsqualität zu finden sind. Konkret sollen folgende Forschungsfragen (FF) untersucht werden:

### **Wirkung des Schüler-Kurzfeedback-Verfahrens**

- FF 1: Lassen sich Hinweise identifizieren bezüglich des Einflusses von Schüler-Kurzfeedback auf a) ausgewählte Skalen der Unterrichtsqualität unter besonderer Berücksichtigung des Schüler-Lehrpersonen-Verhältnisses und b) die Unterrichtsentwicklung und das Unterrichtshandeln der Lehrperson?
- FF 2: Welche Entwicklungen nehmen Schülerinnen und Schüler im Unterricht innerhalb des Schüler-Kurzfeedback-Verfahrens wahr?
- FF 3: Lassen sich Hinweise für Muster identifizieren, die eine positive Wirkung von Schülerfeedback auf die Unterrichtsqualität begünstigen?

### **Praktikabilität des Schüler-Kurzfeedback-Verfahrens**

- FF 4: Wie schätzen Lehrpersonen und Schülerinnen und Schüler die Praktikabilität des Schüler-Kurzfeedback-Verfahrens ein?

Das Hauptziel der vorliegenden Arbeit liegt darin, Hypothesen bezüglich der Wirkung von Schülerfeedback zu entwickeln, die gegebenenfalls in einer größer angelegten, dafür inhaltlich stärker eingegrenzten Nachfolgestudie detailliert untersucht werden könnten.

Zur Klärung der Ausgangslage wird in Kapitel 2 der Wissensstand zur Bedeutung und Funktion von Schülerfeedback beschrieben, dies geschieht unter Berücksichtigung einer Vielzahl von Aspekten, die mit Schülerfeedback im Zusammenhang stehen. Dabei sollen unter anderem auch Variablen herausgeschält werden, die sich aufgrund von Schülerfeedback verändern lassen könnten, beziehungsweise es wird nach theoretischer oder empirischer Evidenz für solche Effekte gesucht.

Auf eine Auswahl dieser Variablen/Aspekte/Effekte wird im Kapitel 3 präziser eingegangen, um ein Schüler-Kurzfeedback-Verfahren zu entwickeln, welches aufschlussreiche Informationen über die ausgewählten Zielintentionen des Unterrichts geben kann und ökonomisch einsetzbar ist. Aus diesen Vorarbeiten leiten sich die bereits erwähnten Forschungsfragen und das Untersuchungsdesign (Kapitel 4) ab. Die Ergebnisse der Untersuchung werden in Kapitel 5 vorgestellt und daraufhin in Kapitel 6 interpretiert, diskutiert und daraus empirisch begründete Hypothesen über die Wirkung und Wirkmechanismen von regelmäßigem Schülerfeedback gebildet. Die Arbeit schließt mit einem praxis- und forschungsbezogenen Ausblick.

## 2 Forschungsstand

### 2.1 Vom Feedback zum Schüler-Feedback: Charakteristik und Definitionen

Obwohl der Begriff Feedback in der Schul- und Arbeitswelt als ein wichtiges Instrument der Qualitätsverbesserung angesehen wird, besteht eine gewisse Unschärfe in der Definition von Feedback. Ditton & Müller (2014) begründen dies vor allem damit, dass der Begriff in sehr unterschiedlichen Forschungs- und Anwendungsfeldern verwendet wird. Die ursprüngliche Definition stammt aus der Kybernetik (Wissenschaft der Steuerung und Regelung von Maschinen, lebenden Organismen und sozialen Organisationen) und bezeichnet in einem Regelkreis einen Messvorgang, der rückgekoppelt wird und darauf ausgerichtet ist, die Differenzen zwischen Soll-Vorgabe und Ist-Zustand aufzuzeigen, um dann entsprechende Korrekturmaßnahmen auszulösen (Landwehr, 2003/2007).

Für Ditton & Müller (2014) erscheint diese Definition aus der Perspektive der Kommunikationsforschung zu eng. Sie verstehen Feedback als „beabsichtigte verbale Mitteilung an eine Person, wie ihr Verhalten oder die Auswirkung ihres Verhaltens wahrgenommen und erlebt worden sind“ (Oberhoff 1978, S.6 zitiert nach Ditton & Müller 2014), dabei besitzt das Feedback keine Allgemeingültigkeit und es steht dem Feedbackempfänger frei, ob er das Feedback für seine künftigen Handlungen berücksichtigt (Ditton & Müller 2014). Landwehr (2003/2007) spricht in diesem Zusammenhang von doppelter Subjektivität.

Obwohl in der Lehr-/Lernforschung Lehrer-Feedback (Lehrperson ist Feedbackgeber und Schülerinnen und Schüler Feedbackempfänger) ebenfalls eine kommunikative Handlung darstellt, ist in diesem Fall eine größere Nähe der Definition zur Kybernetik festzustellen, da hierbei Kommunikation innerhalb eines Steuerungsprozesses zur Erreichung eines Lernziels (Soll-Wert) stattfindet. Demzufolge benutzt Hattie (2007) in seiner Meta-Analyse "Lernen sichtbar machen" die Definition von Sadler (1989) und beschreibt (Lehrer-)Feedback als spezifische aufgaben- oder prozessbezogene Information, die die Lücke zwischen dem was verstanden wurde (Ist-Wert) und dem was es noch zu verstehen gilt (Soll-Wert), schliessen soll. Lehrer-Feedback hat also die Intention einem vorher definierten Soll-Wert näher zu kommen.

Die oben erwähnte doppelte Subjektivität wird im kybernetischen Sinne von Feedback als Störgröße empfunden, welche die Lernsteuerung einschränkt, da die Intention *Annäherung an den Soll-Wert* nicht zwingend oder nicht hinreichend umgesetzt wird. Hattie & Timperley (2007) dehnen deshalb den Feedbackbegriff so weit aus, dass dieser in das „Steuerungsmodul“ des Lernens eingreift. Konkret bedeutet dies, dass Feedback nicht nur auf das Lernziel selbst bezogen wird, sondern unter anderem auch auf den Lernprozess (Was muss getan werden um den

Inhalt zu verstehen?) und die Art und Weise, wie man diesen Lernprozess in Gang hält (Selbstregulation), also neben der kognitiven auch die metakognitive Dimension mitberücksichtigt.

Andere Autoren, wie etwa Butler & Winne (1995), klassifizieren Feedback in ergebnisbezogenes und elaboriertes Feedback. Zweitgenanntes enthält zusätzliche Informationen zur Lösung einer Aufgabe, wie beispielsweise Hinweise auf Lösungsstrategien. Narciss (2006, 2014) plädiert dafür, die unterschiedlichen Formen von elaboriertem und ergebnisbezogenen Feedback zu einem informativ tutoriellen Feedback zu kombinieren. Feedback liefert hier also nicht nur Information eines Soll-Ist-Wert-Vergleichs sondern auch eine Anleitung, wie der Soll-Wert im Falle einer Divergenz erreicht werden kann.

In eine ähnliche Richtung zielt die Methode der formativen Beurteilung, die in ihrer Definition sehr der des Lehrer-Feedbacks ähnelt, allerdings mit einer stärkeren Betonung auf Strategien zur Verringerung der Soll-Ist-Wert-Diskrepanz:

”Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited” (Black & Wiliam 2009, S. 7).

Grob (in Druck) fasst diesbezüglich drei Erklärungs-Ansätze zusammen, wie formative Beurteilung das Lernen unterstützt:

1. Das gegenseitige Einverständnis der Bewertungskriterien zwischen Schülerinnen und Schülern und Lehrpersonen,
2. Die Verbesserung des selbstregulierten Lernens und
3. Das Geben von Feedback auf der Ebene des Lernprozesses

All den beschriebenen Erweiterungen des Feedbackbegriffs ist gemein, dass sie vom Feedbackgeber eine hohe Expertise verlangen. Er muss nicht nur die richtige Lösung oder mögliche Lösungswege der Aufgabe kennen, beziehungsweise eine Vielzahl von Lehr-/Lern-Situationen beurteilen können; er muss auch in der Lage sein, Strategien anzuwenden, die zu einer Annäherung an die Zielvorgaben beziehungsweise zu einer Verbesserung der Lehr-/Lern-Situation führen.

Vor diesem Hintergrund kann man den Prozess des Schüler-Feedbacks ebenfalls als Kommunikationsprozess verstehen, der die Intention beinhaltet eine Annäherung eines Soll-Werts zu erreichen (Verbesserung des Unterrichts). Jedoch erfolgt der Feedbackprozess vom Novizen zum Experten. Gewiss ist bei einem Schüler-Feedback nicht ausgeschlossen, dass auch die Schülerinnen und Schüler geeignete Maßnahmen zur Verbesserung von Unterricht einbringen können, und wie es sich im Verlauf dieser Arbeit zeigen wird, auch sollen. Doch in vielen Fällen ist es die Lehrperson, die entscheidet, welche Anregungen berücksichtigt und welche Maßnahmen ergriffen werden.



Auch wenn Hattie, wie schon in der Einleitung erwähnt, Feedback als einen Prozess sieht, der in beide Richtungen verlaufen soll, so lässt sich feststellen, dass die beiden Richtungen nicht die gleichen Charakteristika aufweisen und somit die oben genannten Definitionen und Methoden vor allem auf Feedbacks zutreffend sind, die der Experte dem Novizen übermittelt. Konsultiert man jedoch die einschlägige Literatur bezüglich Definitionen von Schüler-Feedback, wird man selten fündig und wenn doch, sind diese einerseits wenig ausdifferenziert und weisen andererseits einen verstärkten Fokus auf die methodische Durchführung auf. So beschreiben Herrmann & Höfer (2002, S.29, ff.) Schüler-Feedback als „den systematischen Versuch, die Perspektive von Schülerinnen und Schülern sowie deren Einschätzung des Unterrichts und ihrer eigenen Lernprozesse in die weitere Gestaltung des Unterrichts einzubeziehen“. Bastian, Combe & Lange (2013, S. 90) sehen „Feedback als das methodisch gestützte Miteinander-ins-Gespräch-Kommen und -im-Gespräch-Bleiben und dass diese Gespräche geführt werden, um ‚Verbesserungen des Unterrichts gemeinsam planen und umsetzen‘ (Graf 2001, S. 10) zu können“.

Zwar werden diese Definitionen den Überlegungen bezüglich der kybernetisch/tutoriell/formativen Vorstellung von Feedback, welches eine Annäherung an einen Soll-Wert anstrebt, gerecht, allerdings bleiben die spezifischen Besonderheiten, die bei einem Schülerfeedback auftreten können, unberücksichtigt. Es wird deshalb im Rahmen der vorliegenden Arbeit folgende Arbeitsdefinition vorgeschlagen:

*Schüler-Feedback ist eine prozessbezogene Information von den Schülerinnen und Schülern an die Lehrperson, welche die Qualität des Unterrichts und/oder seiner Zielintentionen aus Schülersicht – vergleichend mit einem subjektiven Referenzwert – beschreibt; dies soll zu einer situativen oder allgemeinen Verhaltensänderung der Lehrperson führen, die wiederum zum Ziel hat, den subjektiven Vorstellungen eines idealen Unterrichts beziehungsweise der vollständigen Erreichung seiner Zielintentionen näher zu kommen.*

Aus dem bisher gesagten wird deutlich, dass bei einem Schüler-Feedback, also einem vertikalen Feedback von unten nach oben, der Einfluss der doppelten Subjektivität eine größere Rolle spielt als bei einem vertikalen Feedback von oben nach unten. Dies kann beispielsweise in der unterschiedlichen Vorstellung oder Einschätzung bezogen auf Unterricht begründet sein (siehe Abschnitt 2.1.2, Reliabilität und Validität von Schülerbeurteilungen) oder seine Ursache in der abweichenden Interpretationen des Feedbacks selbst haben. Der Einfluss personaler Faktoren der Feedbackgeber und Feedbackempfänger ist demzufolge größer. Diese können sowohl auf die Erstellung eines Feedbacks als auch auf die Verarbeitung des selbigen einen Einfluss haben. Da sich im Laufe der Arbeit zeigen wird, dass die Erstellung der Feedbacks zum größten

Teil vorgegeben ist und über die Klasse gemittelt wird, ist vor allem zu klären, inwieweit die personalen Faktoren die Feedbackverarbeitung beeinflussen.

### 2.1.1 Personale Einflussfaktoren auf die Feedbackverarbeitung

Von den personalen Faktoren, die bei der Verarbeitung von Feedback zum Tragen kommen, wurden vor allem Attributionsprozesse, Selbstwirksamkeitsüberzeugungen, Zielorientierungen aber auch Persönlichkeitsmerkmale untersucht. Außerdem können unter anderem Kompetenzen und Kompetenzeinschätzungen bei der Feedbackverarbeitung eine Rolle spielen. Strijbos & Müller (2014) geben hierzu einen aufschlussreichen Überblick. Die für diese Arbeit relevanten Aspekte dieses Beitrags werden in den folgenden Abschnitten zusammengefasst.

#### 2.1.1.1 Einfluss von Attributionsprozessen auf die Feedbackverarbeitung

Unter Attributionen versteht man die Ursachenzuschreibung eines Handlungsergebnisses. Nach Weiner (1985) können diese innerhalb oder außerhalb der eigenen Person liegen, variabel oder stabil und kontrollierbar oder unkontrollierbar sein (siehe Tab. 1).

Tabelle 1: Klassifikationsschema von Attributionen nach Weiner (1985)

	Internal		External	
	stabil	variabel	stabil	variabel
kontrollierbar	Wissen	Lernaufwand	Lernumgebung	Aufgabenwahl
unkontrollierbar	Intelligenz	Gesundheit	Anforderungsniveau	Aufgabenvorgabe

Um eine Handlung zur Annäherung an den Soll-Wert beim Feedbackempfänger zu erreichen, scheint es am günstigsten zu sein, wenn er das Ergebnis der vorausgehenden Handlung internal, variabel und kontrollierbar attribuiert. In dem Fall übernimmt der Feedbackempfänger persönlich die Verantwortung des Ergebnisses und sieht gleichzeitig Anpassungsmöglichkeiten seiner Handlungen, die zu einer Annäherung an die Zielvorgaben führen (Strijbos & Müller 2014).

Dem Feedbackgeber sollte dabei die Aufgabe zufallen, genau diese Attributionen anzuregen und eventuell externale Gründe hinzufügen, um die Akzeptanz eines negativen Feedbacks zu steigern (Strijbos & Müller 2014; Semmer & Jacobshagen, 2010). Allerdings zeigt eine Untersuchung von Lyden et al. (2007), dass weniger die Attributionshinweise der Feedbackgeber die Attribution der Empfänger beeinflusst hat, als vielmehr die Valenz des Feedbacks, also ob es sich um ein positives oder negatives Feedback handelte. Dabei wurden positive Ergebnisse eher mit den eigenen Fähigkeiten (internal) und negative Ergebnisse eher mit der Aufgabenschwierigkeit (external) attribuiert.

Neben der Attribution bezüglich der Handlungsergebnisse können auf Seiten des Feedbackempfängers auch Gründe für das Geben eines Feedbacks zugeschrieben werden. Untersuchungen von Hempel (2008) und Leung und Kollegen (2001) haben gezeigt, dass sich Feedbacks, deren Ursachen in internalen, dispositionalen Aspekten (Sympathie, Emotionen usw.) des Feedbackgebers gesehen werden, negativ auf den Rezeptionsprozess auswirken.

### **2.1.1.2 Einfluss der Selbstwirksamkeitsüberzeugung auf die Feedbackverarbeitung**

Unter Selbstwirksamkeitsüberzeugungen versteht man das eigene Zutrauen, bestimmte Herausforderungen lösen zu können (Bandura 1997). Ist dieses Zutrauen stark ausgeprägt, werden höhere Ziele angestrebt und es wird davon ausgegangen diese Ziele auch erreichen zu können; dabei wird im Falle einer Zielabweichung konstruktiver gehandelt (Bandura 1986).

Bezogen auf den Umgang mit Feedback werden im Falle eines Fehlers laut Srijbos & Müller (2014) günstige Attributionsmuster gezeigt: „Durch eine Ursachenzuschreibung auf eher kontrollierbare Faktoren entsteht Handlungsmotivation [...] und durch externale Attribution wird das eigene Selbst geschützt“ (Bandura & Wood 1989; Lyden et al. 2002; Weiner 1985; Stajkovic & Sommer 2000 zitiert nach Srijbos & Müller 2014, S. 98). Diesbezüglich wiesen Heslin & Latham (2004) nach, dass Feedbackempfänger, die eine hohe Selbstwirksamkeitserwartung besaßen, ihre Leistung verbesserten. Ist dagegen die Selbstwirksamkeitserwartung gering, wird bei negativem Feedback eher auf internale und stabile Faktoren attribuiert (Lyden et al. 2002), was sich schlussendlich in einer geringeren Leistung oder vergleichsweise negativen Leistungsentwicklung widerspiegelt (Bandura 1997).

### **2.1.1.3 Einfluss der Zielorientierung auf die Feedbackverarbeitung**

Zielorientierung beschreibt die Vorstellungen und Überzeugungen, die einer Zielerreichung zugrunde liegen (Srijbos & Müller 2014). Nach Elliot & Dweck (1988) wird dabei die Leistungsorientierung (kompetent erscheinen, Misserfolg vermeiden) von der Lernorientierung (Verbesserung der Kompetenzen, Entwicklung von Fähigkeiten) unterschieden.

In der Verarbeitung von Feedback zeigen Studien, dass die Zielorientierung einen Einfluss auf das Verarbeiten von Feedback hat. Lernorientierte gehen im Falle eines negativen Feedbacks konstruktiver mit diesem um und nehmen es weniger als selbstwertgefährdend wahr (Cianci et al. 2010). Im Falle eines positiven Feedbacks bewirkt die Lernorientierung jedoch eine geringere Leistungssteigerung als bei Leistungsorientierten. Wahrscheinlich sehen Lernorientierte kein weiteres Steigerungspotenzial in der zu bewältigenden Aufgabe und nehmen sie als abgeschlossen wahr (Dweck, 1989). Dagegen reagieren Leistungsorientierte vor allem bei stark ausgeprägten positiven Rückmeldungen „eher mit erhöhten individuellen Zielen und Leistungssteigerung“ (Cianci et al. 2010; Donovan & Hafsteinsson 2006 zitiert nach Srijbos & Müller 2014, S. 100). Erklärt wird dies durch die Befriedigung des, bei Leistungsorientierten, stark aus-

geprägten Bedürfnis nach Kompetenzerleben, was zu einer Motivation für eine weitere Leistungssteigerung führt.

Mit dem letzten Satz könnte der Eindruck erweckt werden, dass es sich bei der Ausprägung der Zielorientierung um ein dispositionales Merkmal handelt. Dies ist aber nur bedingt der Fall. Ci-anci et al. (2010) wiesen nach, dass bei über 99% der Probanden mittels direkter Instruktion, entweder eine Lernzielorientierung oder eine Leistungsorientierung angeregt werden konnte (Zielorientierung hat somit einen dispositionalen und situationalen Charakter). Das bedeutet, dass bei einem Feedback, welches den Lernfortschritt betont, eher eine Lernzielorientierung, und bei einem, das die soziale Bezugsnorm herausstellt, eher eine Leistungsorientierung angeregt wird (Ames 1992; Illgen & Davis 2000). Somit lässt sich durch entsprechende Feedbackformulierung die Zielorientierung beeinflussen, und es bietet sich an, dies situationsspezifisch einzusetzen.

#### **2.1.1.4 Einfluss von Persönlichkeitsmerkmalen auf die Feedbackverarbeitung**

Einige wenige Untersuchungen beschäftigten sich in der Vergangenheit mit dem Einfluss der fünf Persönlichkeitsmerkmale *Neurotizismus*, *Extraversion*, *Offenheit für Erfahrungen*, *Verträglichkeit* und *Gewissenhaftigkeit* (*Big Five*) nach Costa & McCrea (1992). Die Ergebnisse hierzu sind zum Teil uneinheitlich, was sehr wahrscheinlich daran liegt, „dass die fünf Persönlichkeitsdimensionen zwar die Basis für sogenannte *characteristic adaptins* (Hervorhebung im Original) (Einstellungen, Gewohnheiten, Fähigkeiten) darstellen, aber eben auch nur über diese im Zusammenspiel mit Umwelteinflüssen mittelbar und indirekt auf das situative Verhalten wirken“ (McCrea & Costa 2011, zitiert nach Strijbos & Müller 2014, S. 101). Deshalb – und weil die Big Five als relativ stabil gelten und wenig Möglichkeiten bestehen durch entsprechende Anpassungen der Feedbacks auf die Persönlichkeitsmerkmale einzuwirken beziehungsweise diese zu berücksichtigen – wird hier nicht weiter darauf eingegangen.

#### **2.1.1.5 Einfluss von Statusunterschieden zwischen Feedbackgeber und Feedbackempfänger auf die Feedbackverarbeitung**

Gerade in Bezug auf Schülerfeedback ist zu erwarten, dass der Aspekt des Statusunterschiedes im Feedbackprozess eine gewichtige Rolle einnimmt. Allerdings wurden bisher vor allem die Einflüsse des vertikalen Feedbacks von oben (Lehrer, Experte, Vorgesetzter) nach unten (Schüler, Novize, Untergebener) untersucht, und das sowohl im Lehr-Lern- als auch im Unternehmensumfeld (Strijbos & Müller 2014). Die wenigen, bisher durchgeführten Untersuchungen im Lehr-Lern-Kontext, die ein Feedback von unten nach oben thematisieren, werden in Unterkapitel 2.2 ausführlicher dargestellt.

### **2.1.1.6 Einfluss der Sach-Kompetenz des Empfängers auf die Feedbackverarbeitung**

Inwieweit die Sach-Kompetenz des Empfängers für die Verarbeitung von Feedback entscheidend ist, untersuchten unter anderem Kulhavy & Stocks (1989). Sie zeigten, dass die Antwortsicherheit der Probanden die Akzeptanz und Nutzung von Feedback beeinflusst. Je höher die Diskrepanz zwischen der Antwortsicherheit und dem gegebenem Feedback ist (hohe Antwortsicherheit, jedoch falsche Antwort), desto höher ist die Feedbackverarbeitungszeit; zudem tragen in dem Fall komplexere elaborierte Feedbackbedingungen (ähnlich wie beim formativem Beurteilen werden hier Informationen zur richtigen Lösung mitgeliefert) nicht immer zur besseren Wirkung bei (Narciss 2014 nach Mory 2004, Kulhavy & Wager 1993, Kulhavy et al. 1985; Phye & Bender 1989).

### **2.1.1.7 Einfluss der wahrgenommen Kompetenz des Feedbackgebers auf die Feedbackverarbeitung**

Aufgrund der Tatsache, dass Schüler-Feedback ein Feedback vom Novizen zum Experten darstellt, ist tendenziell zu erwarten, dass die Kompetenz des Feedbackgebers vom Feedbackempfänger in Frage gestellt wird. Srijbos et al. 2010 fanden – bezogen auf Peer-Feedbacks im Lehr-Lernkontext – heraus, dass die Feedbacks von hochkompetenten Feedbackgebern am adäquatesten bewertet, diese allerdings auch am ehesten als Bedrohung für das eigene Selbstwertgefühl wahrgenommen wurden.

Im Arbeits- und Organisationskontext wurde der Aspekt der wahrgenommen Kompetenz des Senders breiter untersucht, hierbei wäre zu prüfen, wie weit die gewonnenen Erkenntnisse auch im Lehr-Lernkontext berücksichtigt werden sollten. Die Forschungsarbeiten konzentrierten sich vor allem auf die Glaubwürdigkeit des Feedbacksenders. Diese setzt sich wiederum aus der „Expertise des Feedbacksenders, welche auf Faktoren wie Ausbildung, Erfahrung, Kompetenzniveau und der Vertrautheit mit dem aktuellen Arbeitsgebiet beruht, und der situativen Einschätzung seiner Vertrauenswürdigkeit durch den Empfänger“ zusammen (Griffin 1967 und Ilgen et al. 1979 zitiert nach Srijbos & Müller 2014, S. 110). Steelmann et al. (2004) wiesen nach, dass Feedback von einer Person mit hoher Glaubwürdigkeit im Allgemeinen positiver bewertet wird als von Personen mit niedriger Glaubwürdigkeit. Daneben zeigen Halperin und Kollegen (1976) und Feys und Kollegen (2011), dass die Expertise des Feedbackgebers die Glaubwürdigkeit des Feedbacks und die Bereitschaft sich zu verbessern positiv beeinflusst. Des Weiteren steigert sich dadurch nach Cusella (1982) auch die intrinsische Motivation sich mit den Arbeitsinhalten weiter zu beschäftigen.

Überträgt man diese Erkenntnisse auf das Verfahren Schüler-Feedback, so ist zu befürchten, dass Lehrperson die Glaubwürdigkeit der Schülerinnen und Schüler als nicht besonders hoch einschätzen, da sie sehr wahrscheinlich die Expertise aufgrund geringem Ausbildungsstand, wenig Erfahrung und geringem Kompetenzniveau im Bereich Didaktik/Pädagogik unterdurch-

schnittlich beurteilen. Je nach Situation (Konflikte mit einzelnen oder mehreren Schülerinnen und Schüler) kann auch eine schwach ausgeprägte situative Vertrauenswürdigkeit vorliegen. Somit würden die Feedbacks eher negativ durch die Lehrperson aufgenommen und es läge eine geringe intrinsische Motivation vor, sich z. B. mit der eigenen Unterrichtsentwicklung kritisch auseinanderzusetzen. Um diese Gefahr zu verringern sollte bei der Implementierung eines Schüler-Feedbacks darauf geachtet werden, dass die Schülerinnen und Schüler nicht das Unterrichtshandeln an sich bewerten (was Expertise verlangt) sondern rückmelden, wie der Unterricht bei ihnen ankommt und ob die Zielintentionen (z. B. ob Lernen stattfindet) des Unterrichts erreicht werden (siehe auch Kap. 3).

Die bis dahin dargestellten personalen Faktoren in einem Feedbackprozess vermitteln einen Eindruck der Multikausalität von Feedbackwirkung. Neben den personalen Faktoren nennen Ditton & Arnold (2004) noch drei weitere Gruppen von Faktoren, die auf den *Regelkreis der Unterrichtsverbesserung* einen Einfluss haben. Hierzu zählen die *Merkmale des Feedbacks selbst* (z. B. schriftlich oder mündlich, siehe Kapitel 3), die *Merkmale der auszuführenden Aufgabe* (hier: Anpassung bzw. Verbesserung des Unterrichts) und die *Merkmale der Handlungssituation* (Rahmenbedingungen der auszuführenden Aufgabe).

Aufgrund der vielen Einflüsse, die im Feedbackprozess eine Rolle spielen, kommen die beiden Autoren zu dem Schluss, „dass die Erwartung, dass Rückmeldungen automatisch zu Veränderungen des Handelns führen, naiv wäre“ (Ditton & Arnold 2004, S. 153).

Zusammenfassend sind also zwei Aspekte hervorzuheben. Erstens, Feedback besitzt immer eine vergangenheits- und eine zukunftsorientierte Komponente und zweitens müssen viele Faktoren im Regelkreis Unterrichtsverbesserung zusammenspielen, damit ein Schülerfeedback wirksam wird und somit zu einer Verbesserung von (Physik-) Unterricht führen kann. Einige dieser Faktorengruppen wurden in der jüngsten Vergangenheit teilweise spezifisch für die Situation Schülerfeedback untersucht, zu anderen liegen so gut wie keine Forschungsergebnisse vor. Im folgenden Abschnitt soll der momentane Forschungsstand diesbezüglich dargestellt werden.

### **2.1.2 Reliabilität und Validität von Schülerbeurteilungen**

Feedbackprozesse benötigen zur Ermittlung des Ist-Wertes ein Messinstrument. Im Falle des Schülerfeedbacks wäre dies beispielsweise eine Unterrichtsbeurteilung durch die Schülerinnen und Schüler oder eine Auskunft über den Grad der Erreichung seiner Zielintentionen (z. B. Förderung von Interesse, Entwicklung von Wissen oder kognitive Aktivierung, siehe Kap. 3). Inwieweit ein solches Instrument die Gütekriterien Validität, Reliabilität und Objektivität erfüllt,

wurde in der Vergangenheit mehrfach untersucht (siehe Clausen 2002, Gruehn 2000, Lüdtke et al. 2006, Lüdtke et al. 2007, Marsh 2007 und Wagner 2008).

So bescheinigt beispielsweise Clausen (2002) den Schülerurteilen, aggregiert auf Klassenebene, eine hohe Reliabilität. Lüdtke et al. (2006) bestätigen dies weitgehend (befriedigende Reliabilität mit Ausnahme der Skalen *Pacing* und *anspruchsvolles Üben*), geben aber zu bedenken, dass die Übereinstimmung der Schülerurteile innerhalb der Skalen (hier: *Regelklarheit*, *Zeitverschwendung*, *Schülerpartizipation*, *Individuelle Bezugsnormorientierung*, *anspruchsvolles Üben*, *Pacing*) in vielen Klassen geringe Werte aufweist. Insbesondere bei den Skalen *Regelklarheit* und *anspruchsvolles Üben*. Die Autoren argumentieren, dass nur bei einer ausreichenden Übereinstimmung der Beobachter eine „Aggregation von individuellen Wahrnehmungen zur Erfassung des Klimas in Arbeitsgruppen oder Organisationen“ (Lüdtke et al. 2006, S. 86) sinnvoll ist. Sie stützen sich dabei auf ein Prinzip der Organisationspsychologie: „It is accepted that a necessary precondition for aggregation is a confirmation that there is an agreement among the individuals who form the group with regard to the aggregated construct“ (Cohen, Dovich & Eick 2001, S. 297). Ist dies nicht der Fall, kann die Validität des aggregierten Konstrukts nicht gewährleistet werden (Lüdtke et al. 2006). Die Autoren stellen deshalb einerseits die Frage, ob die Klassenebene tatsächlich die sinnvollste Aggregationsebene darstellt, sehen aber andererseits in einer hohen Varianz den Ansatzpunkt für Maßnahmen, die der Heterogenität innerhalb der Klasse Rechnung tragen.

Grundsätzlich stellt sich bei der Beurteilung von Unterricht die Frage, ob es einen *objektiv idealen Unterricht* wirklich gibt, auf den sich mögliche Beurteilungsskalen beziehen können. Insbesondere Clausen (2002) nimmt sich der Frage an, und stellt die Lehrpersonen-, Beobachter- und Schülerperspektive einander gegenüber. Er argumentiert, dass sich Unterricht (unabhängig der Beobachterperspektive) immer nur indirekt, über die Wahrnehmung der Teilnehmenden, beziehungsweise Beobachter, erfassen lässt. „Keine der drei Sichtweisen liegt näher an der Unterrichtswirklichkeit als die anderen Perspektiven“ (Clausen 2002, S. 186). Dies zeigt sich in den geringen gemeinsamen Varianzanteilen der Urteile aus den verschiedenen Perspektiven. Er betont dabei die Ähnlichkeit zu Befunden aus der psychologischen Forschung, in denen die Perspektive von sozialen Wahrnehmungen verglichen werden und fordert, die Varianz zwischen den Gruppen nicht als Fehler sondern als Ausdruck von *perspektivenspezifischer Validität* zu betrachten. Zielspezifisch solle man auf die entsprechenden Daten zurückzugreifen. So empfiehlt Clausen (2002) für die Vorhersage von kognitiven oder psychosozialen Entwicklungen das Heranziehen von Schülerbeurteilungen, während bei einer differenzierten Bewertung des Unterrichtsgeschehens der außenstehende Beobachter zu bevorzugen ist.

Der Begriff *perspektivenspezifische Validität* widerspricht sich bei genauer Betrachtungsweise, eignet sich jedoch gut, um die subjektive „Wahrheit“ der Beobachter zu beschreiben. Auch hier bietet sich an, ähnlich wie bei der Varianz der Schülerurteile (s.o.), die Diskrepanzen in der

Wahrnehmung zwischen Schülerinnen und Schüler und Lehrperson für die Unterrichtsentwicklung zu nutzen.

Allerdings können Unterschiede zwischen den Perspektiven nicht über alle Altersgruppen und Merkmale des Unterrichts mit der *perspektivenspezifischen Validität* begründet werden. So zeigt Lenske (2016, S. 311) in ihrer Mixed-Method-Untersuchung zur Konstruktvalidität von Schülerfeedback aus der Perspektive von Grundschulern, dass diese „Facetten einzelner Merkmale nicht hinreichend valide beurteilen konnten“. Dies traf auf „einzelne Aspekte der Schülerorientierung (z.B. das Einbringen eigener Ideen), der Aktivierung (z. B. permanente sinnvolle Beschäftigung), der Klarheit (z. B. das Verwenden guter Beispiele) oder der Strukturierung (z. B. reibungslose Übergänge)“ zu. Hingegen „wurden ausgewählte Aspekte bezüglich der Klassenführung, des lernförderlichen Klimas, der Motivierung, der Klarheit und Strukturiertheit und der Aktivierung kompetent beurteilt“ (ebd., S. 310). Die Autorin bezeichnet das Feedback von Grundschulern demzufolge – und aufgrund der Überlegung, dass die Nutzung des Lernangebots durch die Schülerinnen und Schüler vor allem dann gegeben ist, wenn die Schülerinnen und Schüler das Unterrichtsangebot als gut und förderlich wahrnehmen (Hattie 2013) – als aussagekräftig.

Diese Überlegung führt zum Aspekt der Kriteriumsvalidität, genauer zur prädiktiven Validität von Schülerfeedback. Hier zeigt sich, dass Schülerfeedback eine gute Voraussagekraft bezüglich der Leistungsentwicklung besitzt. Gruehn (2000) hält fest, dass sich 7 bis 20% der Varianz von Lernparametern über die von Schülern beurteilten Unterrichtsmerkmale erklären lässt. Helmke, Schneider & Weinert (1986) konnten 16% der Leistungsunterschiede zwischen Schulklassen mit den von Schülerinnen und Schüler wahrgenommenen Unterrichtsvariablen erklären (zitiert nach Clausen 2002). Dies wird in der Studie von Clausen (2002) bestätigt und er weist der Schülersichtweise eine höhere prädiktive Validität nach als die der Beobachter- und vor allem der Lehrersichtweise.

#### **2.1.2.1 Mögliche Verzerrungen von Schülerwahrnehmungen**

Trotz der geschilderten Evidenz für die Verlässlichkeit und Validität von Schülerurteilen ist mit Verzerrungen der Urteile von Schülerinnen und Schüler zu rechnen. So weist Clausen (2002) auf die beschränkte Generalisierbarkeit der oben erwähnten Ergebnisse hin, indem er anfügt, dass globale Urteile der Schülerinnen und Schüler bezüglich einer Lehrperson einen erhöhten Einfluss auf die Differenziertheit der Beurteilung von Unterrichtsqualitätsmerkmalen haben. Dies bedeutet, dass beispielsweise eine Lehrperson, die von den Schülerinnen und Schüler insgesamt als *schlecht* angesehen wird, auch in allen Einzelmerkmalen (z. B. „Klarheit der Lehrperson“) als eher *schlecht* beurteilt wird, auch wenn sie das nicht ist. Diese Form der verzerrten Wahrnehmung wurde auch in anderen Studien beschrieben (vgl. Haladyna & Hess 1994) und



als Halo-Effekt bezeichnet. Es existieren zur Erklärung des Effekts drei Modelle (vgl. Lance et al. 1994):

- *salient dimension halo*: Ein hervorstechendes Einzelmerkmal überlagert die Beurteilung aller anderen Merkmale,
- *general impression halo*: Der Gesamteindruck der Merkmale wirkt auf die Beurteilung der Einzelmerkmale zurück,
- *inadequate discrimination halo*: Die Fähigkeit zur Trennung der Einzelmerkmale ist unzureichend.

Nach Lance et al. (1994) beschreibt das Modell des *general impression halo* das Phänomen am besten, dabei spielt es keine Rolle in welchem Kontext die Beurteilung stattfindet. Wagner (2008) wies nach, dass dieser Effekt durch Aggregation auf Klassenebene wahrscheinlich nicht eliminiert werden kann, da zwischen den Klassen (gleiches Fach, gleiche Lehrperson) das Globalurteil bezüglich einer Lehrperson stärker korreliert als die differenzierte Beurteilung von Unterrichtsmerkmalen.

Eine weitere Form der Urteilsverzerrung, die insbesondere bei Schülerurteilen auftreten kann, ist der sogenannte *Dr.-Fox-Effekt*. Dieser beschreibt die Möglichkeit, dass durch charismatisches Auftreten der Lehrperson ein Gefühl des Lernfortschritts bei Schülerinnen und Schülern erzeugt wird, der in Wirklichkeit nicht stattfindet (Hofer 1981 zitiert nach Ditton 2002). Allerdings wurde der Effekt bisher nur bei kurzzeitigen Interventionen nachgewiesen und bezog sich auf Einzelaspekte. Da sich nach Gerstenmaier (1975) Schülerurteile auf Erwartungsmuster aufbauen, die durch schulische Sozialisation geprägt werden und nicht individuell-persönliche Präferenzen oder Vorurteile einzelner Schüler widerspiegeln, erscheint die These, dass sich Schülerinnen und Schüler durch charismatisches Auftreten auf längere Sicht beeinflussen lassen, als unwahrscheinlich.

Außerdem gilt die *selbstdienliche Verzerrung* in Bezug auf eine *soziale Erwünschtheit* als weiterer typischer Urteilsfehler. Unter *selbstdienlicher Verzerrung* versteht man die Tendenz, das eigene Verhalten beziehungsweise die eigenen Fähigkeiten besser zu beurteilen als sie der Wirklichkeit entsprechen (Lenske 2016). Würde ein Schülerfeedback zum großen Teil eine Selbstauskunft darstellen (z. B. „Habe ich im Unterricht mitgedacht?“), könnte dieser Bias an Bedeutung gewinnen (abhängig vom Attributionsstil der Schülerinnen und Schüler). Je größer dabei die Unerwünschtheit oder Erwünschtheit des zu beurteilenden Merkmals ist, desto höher ist die *selbstdienliche Verzerrung* (John & Robins 1993, zitiert nach Clausen 2002). Deshalb scheint trotz allem die Verzerrung, die durch Schüler-Feedback entsteht, geringer zu sein als jene durch Selbstevaluation von Lehrpersonen. Lehrpersonen sehen in der Selbstevaluation des Unterrichts oft auch eine Selbstbeurteilung (vgl. u.a. Hook & Rosenshine 1979 und Koziol & Burns 1986, zitiert nach Clausen 2002, siehe auch Abschnitt 2.1.4). Und da Unterrichtserfolg gesellschaftlich erwünscht ist und das Studiendesign sichtbar die Handlung der Lehrperson in

den Fokus stellt, unterläge dieser auch einer stärkeren selbstdienlichen Verzerrung (vgl. Wagner 2008).

Je nach zeitlichem Verlauf eines Feedbackprozesses, kann auch eine Einschränkung der Objektivität aufgrund von *Kommunikation unter den Beurteilenden* auftreten (Kenny 2004), dies vor allem, wenn der Zeitraum zwischen mehreren Feedbacks kurz genug ist, um sich an Aussagen, die eventuell nach der Feedbackerstellung von Mitschülern getätigt wurden, zu erinnern und diese in das darauffolgende Feedback einfließen zu lassen.

Des Weiteren hat Wagner (2008, S.132 ff.) Einflüsse auf die Wahrnehmung von Unterricht durch Schülerinnen und Schüler identifiziert:

*Innerhalb der Klasse* werden [...] Unterrichtsmerkmale...

1. ...von Mädchen meist positiver als von Jungen beurteilt.
2. ... von Schülerinnen und Schüler mit höheren Testleistungszuwächsen (insbesondere im jeweiligen Fach) in der Regel besser beurteilt.
3. ...zumeist positiver eingeschätzt, wenn bessere Noten im jeweiligen Fach gegeben werden, während die jeweilige Fachnote am Ende des vorangegangenen Schuljahres (die allerdings z. T. auch nicht von der anderen Lehrkraft stammt) kaum eine Rolle spielt.
4. ...tendenziell weniger günstig eingeschätzt, wenn „bessere“ Noten im jeweils anderen Fach am Ende des Schuljahres erwartet werden.

*Auf Klassenebene* werden [...] Unterrichtsmerkmale...

1. [...]
2. ...mit zunehmend „besseren“ Noten am Ende des Schuljahres teilweise positiver beurteilt, während auch hier – analog zu den Ergebnissen auf der Ebene innerhalb von Klassen – die Noten am Ende des vorangegangenen Schuljahres keine Rolle spielt.

Insbesondere der zweite Punkt auf Klassenebene lässt, laut Autor, die vorsichtige Interpretation zu, dass Schülerinnen und Schüler Unterricht positiver beurteilen, wenn die Lehrperson, bezogen auf den tatsächlichen Leistungszuwachs, *zu gute* Noten vergibt.

Die Einflüsse durch Verzerrungen auf die Ergebnisse von Schülerurteilen sind demnach mannigfaltig, jedoch folgert Gruehn (2000) auf Grundlage ihrer Daten, dass diese Beeinflussung durch subjektive Beurteilungstendenzen oder sympathiebedingte Parteilichkeiten insgesamt gering ausfällt.

Insgesamt lässt sich festhalten, dass Schülerbeurteilungen, aggregiert auf Klassenebene, als objektiv, reliabel und prädiktiv valide gelten. Obwohl diese den üblichen – durch die besondere Situation „Unterricht“ z. T. auch verstärkten – Verzerrungen unterliegen, die bei subjektiven Wahrnehmungen einer sozialen Situation entstehen. Befürchtungen, Schülerinnen und Schüler würden bei Beurteilungen von Unterricht ihre Bequemlichkeit als Maßstab benutzen oder diese als Rachewerkzeug missbrauchen, werden von der Unterrichtsforschung weitestgehend widerlegt. "Schüler sind also - mit ihrer spezifischen Perspektive - kompetente Beurteiler von Unterricht und es bietet sich an, ja es drängt sich geradezu auf, ihre Einschätzungen für die Verbesserung von Unterricht zu nutzen" (Ditton & Arnoldt 2004, S. 168). Nichtsdestotrotz weichen Schülerurteile bezüglich guten Unterrichts, trotz ihrer Qualität, von den Beurteilungen der Lehr-

personen und Fremdbeobachter ab. Es wird daher gefordert, unterschiedliche Perspektiven bei der Beurteilung von Unterrichtsqualität miteinzubeziehen (Clausen 2002).

Schülerfeedback kommt dieser Forderung, zumindest in großen Teilen, nach, indem die Lehrerperspektive um die Schülerperspektive erweitert wird. Um einen Abgleich der Perspektiven durchzuführen ist es jedoch erforderlich die Ergebnisse der Schülerbeurteilung in der Klasse transparent zu machen und, wenn nötig und möglich, zu diskutieren. Auf diese Weise könnten Differenzen zwischen den verschiedenen Perspektiven aufgeklärt und weitere Maßnahmen besprochen werden.

Die möglichen Verzerrungen, wie beispielsweise der *Halo-Effekt*, könnten abgemildert werden, indem nicht der Unterricht oder die Lehrperson beurteilt wird, sondern die für den Lernprozess entscheidenden Empfindungen und Aktivitäten beschrieben und skaliert werden. Dies kann über eine Einstufung der empfundenen Rahmenbedingungen (z. B. Ich empfand die Atmosphäre des Unterrichts für das Lernen hilfreich.) oder des eigenen Lernverhaltens geschehen (z. B. Der Unterricht hat mich angeregt genauer über den Inhalt nachzudenken). Hierbei könnte allerdings der Effekt der selbstdienlichen Verzerrung (Der/die Lernende beschreibt eine höhere kognitive Aktivität, weil diese gesellschaftlich erwünscht ist) stärker in Erscheinung treten. Aus den oben und in Abschnitt 2.1.4 genannten Gründen ist allerdings zu erwarten, dass diese geringer ausfällt als bei einer Selbstbeurteilung durch die Lehrperson.

Außerdem sollte das Ausmaß der Streuung bei der Interpretation der Schülerbeurteilungen berücksichtigt werden. Die Lehrperson erhält dadurch Hinweise, ob von ihrem Unterricht alle oder nur gewisse Teile der Klasse profitieren.

### **2.1.2.2 Mögliche Ursachen für die Güte der Schülerbeurteilung**

Gründe für die zum Teil überlegene prädiktive Validität von Schülerbeurteilungen, insbesondere gegenüber der Selbstwahrnehmung der Lehrperson, könnten in der Langzeiterfahrung liegen, die Schülerinnen und Schüler mit einer Lehrkraft und ihrem Unterricht besitzen (Clausen 2002, Ditton 2002). So fällt es externen Beobachtern schwerer zu beurteilen, ob beispielsweise ein problematischer Unterrichtsverlauf einer unruhigen Klasse oder einer unzureichenden Kompetenz der Lehrperson zuzuschreiben ist. Ähnlich verhält es sich mit der Selbstwahrnehmung der Lehrperson, die selten in die Lage versetzt wird, den Unterricht derselben Klasse bei einer anderen Lehrperson mitzerleben und mit dem eigenen Unterricht zu vergleichen.

Außerdem ist es für die Validität einer Messung entscheidend, ob tatsächlich unter den natürlichen Rahmenbedingungen gemessen wurde. Diese Frage stellt sich häufig bei Unterrichtsstunden, die von einem externen Beobachter (z. B. Videostudien) analysiert werden. Denn obwohl Lehrpersonen die eigene videographierte Unterrichtsstunde meist als repräsentativ einschätzen (vgl. Seidel et al. 2003), kann nicht ausgeschlossen werden, dass Fremdbeobachtungen zu

einer Verfälschung des alltäglichen Unterrichts führen können. Wohingegen Schüler-Feedback in das normale Unterrichtsgeschehen eingebettet ist und somit die äußeren Einflussfaktoren eliminiert werden und der Grad der Verfälschung gesenkt wird (Clausen 2002).

Die wahrscheinlich gewichtigste Rolle spielt jedoch, dass die Schülerinnen und Schüler als Adressaten des Unterrichts am besten beurteilen können, wie der Unterricht auf sie wirkt, beziehungsweise was er bei ihnen bewirkt. Als externer Beobachter oder als Lehrperson ist es beispielsweise schwierig, Indikatoren für die kognitive Aktivierung beziehungsweise Passung oder Motivation zu finden. "Ein Schüler weiß in der Regel jedoch besser, ob ihn das Thema oder die Aufgabe wirklich interessiert hat und ob er *on task* (Hervorhebung im Original) war oder eventuell nur so getan hat" (Lenske 2016, S. 80).

### **2.1.3 Weitere mögliche Vorteile von Schülerfeedback**

Lenske (2016) sieht in dem Verfahren Schülerfeedback neben seiner Güte noch weitere Vorteile für die Unterrichtspraxis. Diese liegen in der *Ökonomie* des Verfahrens, der mit dem Schülerfeedback einhergehenden *Schülerorientierung* und in der *Bedeutsamkeit* des Urteils für das Verhalten der Schüler selbst.

Die *Ökonomie* des Verfahrens bezieht sich vor allem auf den Vergleich zum Alternativverfahren Fremdbeobachtung. Es sind keine Absprachen oder Stundenplanänderungen nötig und der finanzielle Aufwand hält sich ebenfalls in Grenzen. In der Praxis kann allerdings eine Auswertung eines oder mehrerer Klassensätze Schülerdaten eine Mehrarbeit bedeuten, die manche Lehrperson abschrecken könnte.

*Schülerorientierung* beschreibt, inwieweit Schülerinnen und Schüler in die Unterrichtsentwicklung miteinbezogen werden. Ein Schülerfeedback kann einen wesentlichen Beitrag zur Unterrichtsentwicklung leisten. Werden Schülerinnen und Schüler in die Unterrichtsentwicklung mit einbezogen sinkt der Grad des Vorbestimmten und das Autonomieempfinden steigt, was sich wiederum positiv auf die Motivation der Schülerinnen und Schüler auswirken kann.

Der Aspekt der *Bedeutsamkeit* wurde bereits zum Ende des Abschnitts 2.1.2 in einem Einschub kurz angedeutet. Da die Schülerinnen und Schüler Adressaten von Unterricht sind, ist es für sie besonders bedeutsam, ob sie den Unterricht als gut beziehungsweise schlecht empfinden. Entscheidend - und unter Berufung auf die Motivationstheorie leicht begründbar - ist nicht, ob der Unterricht nach objektiven Kriterien wirklich gut ist, sondern ob die Schülerinnen und Schüler ihn als gut wahrnehmen. Lehrerhandeln ist demnach besonders effektiv, wenn die Schülerinnen und Schüler es als förderlich wahrnehmen (Hattie 2013).

### 2.1.4 Selbstwahrnehmung des eigenen Unterrichts

Die Beurteilung des eigenen Unterrichts durch die Lehrperson gehört in seiner unsystematischen Form zum Alltagsgeschäft. Eine solche Selbstevaluation kann während des Unterrichts durchgeführt werden, kurz danach oder bei der erneuten Vorbereitung eines bereits unterrichteten Inhalts.

Wenn Lehrpersonen über eine entsprechende Berufserfahrung verfügen und eine Klasse bereits längere Zeit unterrichten, scheinen sie meist in der Lage, Anhaltspunkte zu identifizieren, die ihnen Aufschluss darüber geben, ob die Ursachen für das Scheitern einer didaktischen Herangehensweise eher in der Klasse oder in der mangelhaften Durchführung zu suchen sind. Solche Ursachenzuschreibungen unterliegen jedoch auch dem jeweiligen Attributionsstil der Lehrperson.

Wie in Abschnitt 2.1.2.1 bereits angedeutet, wird „in anglo-amerikanischen Studien zur Teacher-Effectiveness die Beurteilung des Unterrichts durch den Lehrer häufig als Selbstbeurteilung verstanden“ (vgl. u.a. Hook & Rosenshine 1979 und Koziol & Burns 1986, zitiert nach Clausen 2002). Wenn dies zutrifft, ist zu erwarten, dass systematische Fehler, wie die selbstdienliche Verzerrung, eine gewichtige Rolle spielen. Clausen (2002) findet hierzu Anhaltspunkte, indem er, im Vergleich zur Schülersicht, geringere Zusammenhänge zwischen der Selbstbeurteilung und der Unterrichtsqualität belegt. Vor allem in Bezug auf die Effektivität von Unterricht. Der Autor schränkt allerdings ein, dass ohne ein echtes Validitätskriterium, eine selbstdienliche Verzerrung nicht einwandfrei zu belegen ist.

Ungeachtet dessen besitzt die Lehrpersonenbeurteilung, wie oben angeführt, im Vergleich zu den beiden anderen Perspektiven, die geringste Erklärungskraft künftiger schulischer Entwicklungsverläufe (Clausen 2002). Obwohl Schülerbeurteilungen demnach nicht nur die Perspektive auf die eigene Unterrichtsqualität erweitern, sondern nach aktuellem Forschungsstand auch die Objektivität und die prädiktive Validität der Unterrichtsbeurteilung erhöhen, ist in der Praxis die mehr oder weniger systematische Selbstbeurteilung von Lehrpersonen häufig das einzige Instrument, das die Unterrichtsqualität direkt misst und in die weitere Unterrichtsentwicklung miteinbezogen wird.

Doch selbst wenn das Verfahren Schüler-Feedback durchgeführt wird (wie beispielsweise im Hochschulbereich, siehe Abschnitt 2.2.1),

„führt das Feedback, das den Lehrpersonen gegeben wird, selten zu Verbesserungen in ihrem Unterrichtsstil oder der Effektivität der Kurse. [...] Die mangelnde Nutzung von entsprechenden Bewertungen durch Lernende aus der Grundschule und aus den High-Schools sollte uns zu denken geben. [...] Lernen muss von den Lehrpersonen aus der Perspektive der Lernenden betrachtet werden, damit sie besser verstehen, wie das Lernen aus der Sicht der Lernenden aussieht und wie es sich für sie anfühlt.“ (Hattie 2013, S. 138-139).

Dieser seltene Einsatz von Schülerfeedback im Regelschulalltag spiegelt sich auch in der Unterrichtsforschung wider. Obwohl das Messinstrument Unterrichtsbeurteilung durch Schülerin-

nen und Schüler die notwendigen Gütekriterien erfüllt, gibt es auf der Stufe der Regelschule nur wenige Forschungsarbeiten, die deren Wirkung in einem Feedbackprozess untersuchen. Die aktuellsten Untersuchungen zu dieser Thematik werden im folgenden Unterkapitel vorgestellt.

## **2.2 Wirkungen von Schülerfeedback**

Aufgrund des Mangels an Untersuchungen im Regelschulbereich soll zunächst auf Untersuchungen im Hochschulbereich zurückgegriffen werden, um dann eine Brücke zu den Untersuchungen im Regelschulbereich zu schlagen.

### **2.2.1 Ausgewählte Untersuchungen an Hochschulen**

Da in deutschsprachigen Hochschulen das Verfahren der Lehrevaluation bereits seit den 1970er Jahren regelmäßig Anwendung findet (Diehl, 2001), ist dessen Wirkung auf dieser Bildungsstufe relativ gut untersucht. International wurden sogar seit den 1960er-Jahren Untersuchungen durchgeführt, welche die Effekte von evaluierten Lehrveranstaltungen überprüften (Lang & Kersting, 2007). Cohen (1980) und L'Hommedieu, Menges & Brinko (1990) führten aufgrund der Datenlage Metaanalysen durch. Dabei bezog sich der Erhebungszeitraum auf die Dauer eines Semesters, währenddessen in einem Kontrollgruppendesign in der Mitte (Experimentalgruppe) und zum Ende (Experimental- und Kontrollgruppe) des Semesters eine Evaluation durchgeführt wurde. Die Effektstärken (Cohens  $d$ ) der Experimentalgruppen lagen zwischen  $d = 0,38$  (Cohen, 1980) und  $d = 0,34$  (L'Hommedieu, Menges & Brinko, 1990).

Dabei zeigte sich, dass sich die Effekte vergrößerten, wenn die Dozierenden zusätzlich beraten wurden (Cohen, 1980:  $d = 0,20$  für die reine Feedbackbedingung und  $d = 0,64$  für die Bedingung Feedback mit Beratung; entsprechend L'Hommedieu, Menges & Brinko  $d = 0,18$  und  $d = 0,55$ ) (Lang & Kersting, 2007). Eine Evaluation durchzuführen, ohne weiteren formativen Input, scheint also nur geringe Effekte zu erzielen.

Der Erfolg eines Feedbackverfahrens könnte demnach davon abhängig sein, inwieweit sich Dritte an der Interpretation, Beratung und den Folgen der Ergebnisse beteiligen. Rindermann (2009) unterscheidet sechs Invasionsstufen im Umgang mit Lehrveranstaltungsevaluation (LVE) (vgl. Schmidt, 2008):

1. Sensibilisierungs-Hypothese: Erhöhung der Aufmerksamkeit der Lehrenden für das eigene Lehrverhalten, was zu einer Verbesserung des selbigen führen soll.
2. Feedback-Hypothese: Anpassung des Lehrverhaltens auf Basis von Informationen aus Sicht der Lernenden für die Lehrenden. Dies soll den Ist-Wert an den Soll-Wert annähern.

3. Diskursmodell: Förderung des Dialogs mit Teilnehmenden durch LVE, was zu einer Konkretisierung der schriftlichen Evaluation und zu einer verstärkten Orientierung an den Interessen der Lernenden und somit zu einer Verbesserung der Lehrveranstaltung führen soll.
4. Beratungsansatz: Förderung der Metakompetenz der Lehrenden durch Beratung von Experten auf Grundlage der LVE-Ergebnisse zur Steuerung und somit auch zur Verbesserung des Lehrverhaltens.
5. Kompetenzentwicklungsmodell: Entwicklung und Verbesserung der lehrbezogenen Kompetenzen durch Workshops/Trainings, die an die LVE anschließen.
6. Institutionelle Verankerung: Stärkung des Interesses an guter Lehre aufgrund finanzieller/vertraglicher Koppelung an die LVE-Ergebnisse.

Die Punkte 4-6 unterscheiden sich von den Punkten 1-3 dahingehend, dass diese nur mit einem erhöhten Personal- und Finanzierungsaufwand zu bewerkstelligen sind. Daher scheinen die Zahlen (beruhend auf freiwilliger Selbstauskunft), die in einer Studie im Auftrag der Deutschen Hochschulrektorenkonferenz (2010) ermittelt wurden, nicht zu überraschen: Zwar wurden in 89 % der Fälle die Ergebnisse innerhalb des Instituts/Fachbereichs erörtert, aber nur bei 50 % wurden von der Leitung Empfehlungen ausgesprochen und in 30,2 % der Fälle Zielvereinbarungen getroffen. Allerdings wurden auch nur in 32,7 % der Fälle die Ergebnisse den individuell Beteiligten zugänglich gemacht, was die Grundlage für Stufe 3 darstellen würde.

Schmidt (2008) macht deshalb unter anderem eine mangelnde Ausschöpfung der LVE-Verfahren für deren geringen Erfolg verantwortlich. Da es für die Wirksamkeit einer Evaluation von Bedeutung ist, inwieweit diese in weitere Maßnahmen eingebettet ist, könnte das Fehlen von geeigneten Maßnahmen die Wirkungskette von Evaluation unterbrechen. Der Autor nennt diesbezüglich vier Punkte:

- (1) [...] die LVE mit persönlich bedeutsamen Effekten für die Lehrenden (z. B. direkter Dialog, Feedback von Stärken und Schwächen, finanzielle Kompensation) verknüpfen. (2) [...] Folge- und Anschlussmaßnahmen an LVE konstruieren (z. B. Dialog mit Studierenden; Beratungsgespräch; Workshops) [...] (2.3) [...] eine Einbettung der LVE in ein Rahmenkonzept (z. B. Verknüpfung mit anderen Instrumenten; selbstverständlicher Teil professioneller Entwicklung; Bezug zu anderen Elementen der leistungsbezogenen Entlohnung). [...] (2.4) [...] Veröffentlichung der LVE-Ergebnisse (z. B. Dialog mit Studierenden; institutionelle Verankerung). (Schmidt 2008, S. 14)

Allerdings zeigt Rindermann (2009) in seinen, von 1993 bis 1996 durchgeführten Untersuchungen zum hochschuldidaktischen Diskursmodell (Evaluation mit integrierter Studenten-Dozentenbesprechung), dass bei einem Prä- (Semestermitte) Postvergleich (Semesterende) keine positiven Effekte festzustellen sind. Zu ähnlichen Ergebnissen kommt Gediga (2000), der bei einigen Items im Prä-/Postvergleich sogar Verschlechterungen ermittelt hat. Dieses Bild änderte sich allerdings als die Studierenden gefragt wurden, ob sie Veränderungen aufgrund des Verfahrens bemerkt hätten. In dem Fall gaben 47,5 % an, dass sich die Dozierenden „ziem-

lich“ oder „sehr bemüht“ hätten, kritische Punkte zu verbessern. 46,9 % der Befragten nahmen eine „mittelmäßig“ bis „ziemliche“ und 13,3 % „wenig“ Verbesserung wahr.

Diese Widersprüchlichkeit könnte einerseits damit zu erklären sein, dass die geschlossenen Items der Prä-/Posterhebung ein zu grobes Instrument darstellen, um leichte Verbesserungen abzubilden und andererseits könnte sich aufgrund der Evaluation ein höherer Anspruch bei den Studierenden gegenüber der Lehrveranstaltung und des/der Dozierenden entwickelt haben, der sich in einer strengerem Bewertung der/des selbigen widerspiegelt. Es ist also nicht auszuschließen, dass eine Evaluation mit anschließender Besprechung und Diskussion über die Ergebnisse eine Verbesserung der Lehre hervorruft.

Die hier vorgestellten Ergebnisse aus dem Hochschulbereich geben zwar Anhaltspunkte, wie ein erfolgreiches Schüler-Feedback im Regelschulbereich durchgeführt werden könnte, eins-zu-eins lassen sich die Ergebnisse jedoch kaum übertragen. Obwohl die Zielintention die gleiche ist (nämlich Entwicklung der Lehre), so sind doch die Rahmenbedingungen deutlich verschieden (z. B. Klassengröße, Menge der Unterrichtszeit, fester Klassenverbund). Der offensichtlichste Unterschied liegt im jüngeren Alter der Schülerinnen und Schüler. Aber auch der Professionalisierungsgrad bezüglich Lehre ist bei Lehrpersonen im Regelschulbereich meist höher, da diese eine didaktisch/pädagogische Ausbildung mit hohem mentorisiertem Praxisanteil genossen haben. Unter Hochschuldozierenden ist ein vergleichbarer didaktisch/pädagogischer Ausbildungsgrad deutlich seltener anzutreffen. Bezogen auf ein Evaluationsverfahren lässt sich deshalb vermuten, dass ein damit verbundenes Betreuungsangebot durch fachdidaktische Experten an Hochschulen größere Erfolge zeitigt als an Regelschulen.

### **2.2.2 Ausgewählte Untersuchungen im Regelschulbereich**

Seit Hatties (2013) Metaanalyse „Lernen sichtbar machen“ wurde vermehrt gefordert, die Sichtweise der Schülerinnen und Schüler auf das Unterrichtsgeschehen stärker zu berücksichtigen und sie somit in die Unterrichtsentwicklungsprozesse miteinzubeziehen. So entstanden mehrere Onlineportale (z. B. EMU der Universität Koblenz-Landau, LiU der Universität Dortmund, SEfU der Siegfried-Schiller-Universität Jena, SEP des Landes Brandenburg), die es Lehrpersonen möglich machen, auf Instrumente der Unterrichtsevaluation zuzugreifen. Außerdem wurden vermehrt Beiträge und Bücher publiziert, die unterschiedliche Vorschläge zur Durchführung von Schülerfeedbacks enthalten (z. B. Berger 2013, Schneider 2015). Im Bundesland Bayern soll gar die Einholung von Schülerurteilen während des Referendariats verpflichtend werden (Zöller 2016).

Bezüglich der empirisch belegten Wirkung von Schüler-Feedback auf beispielsweise die Unterrichtsentwicklung oder die Unterrichtsqualität im Regelschulbereich ist der Kenntnisstand allerdings begrenzt. Die Forschungsarbeiten beschränken sich häufig auf praxisbezoge-



ne/beschreibende Darstellungen (z. B. Bastian, Corbe & Langer, 2001; Kretschmer 2013; Eikenbusch 2001). Einige der wenigen Arbeiten, welche die Wirkung von Schülerfeedback systematisch erforschten, sind jene von Ditton & Arnold (2004) und Gärtner & Vogt (2013). Da die beiden Arbeiten den Ausgangspunkt für die vorliegende Forschungsarbeit darstellen, werden sie im Folgenden etwas ausführlicher beschrieben.

Bei ersterer Forschungsarbeit handelt es sich um eine Längsschnittuntersuchung, deren Zeitraum ein Schuljahr umfasste, wobei die erste Erhebung im November und die zweite zu Schuljahresende stattfand. Es wurden zu beiden Zeitpunkten sowohl die Schülerinnen und Schüler als auch die Lehrpersonen mit Hilfe eines Fragebogeninstruments befragt. Während die Schülerinnen und Schüler über die zentralen Dimensionen des Unterrichts (QuAIT-Modell - Qualität, Angemessenheit, Anregung, Zeitnutzung und soziale Aspekte; nach Slavin 1996, siehe Abschnitt 4.1.1) befragt wurden, forderte man die Lehrpersonen auf, über drei unterschiedliche Aspekte Auskunft zu geben: Erstens sollten sie die Klassen bezüglich Selbständigkeit, Ausdauer, Aufmerksamkeit, Disziplin, Leistung und Interesse einschätzen; zweitens versuchen, die Beurteilung durch Schülerinnen und Schüler zu prognostizieren und drittens die Schüler-Rückmeldungen auf einer Metaebene bewerten. Bei Letzterem ging es unter anderem darum, inwiefern sie die Schülerinnen und Schüler als kompetente Rückmelder einschätzten und ob sie die Rückmeldungen für den Unterricht für verwertbar hielten.

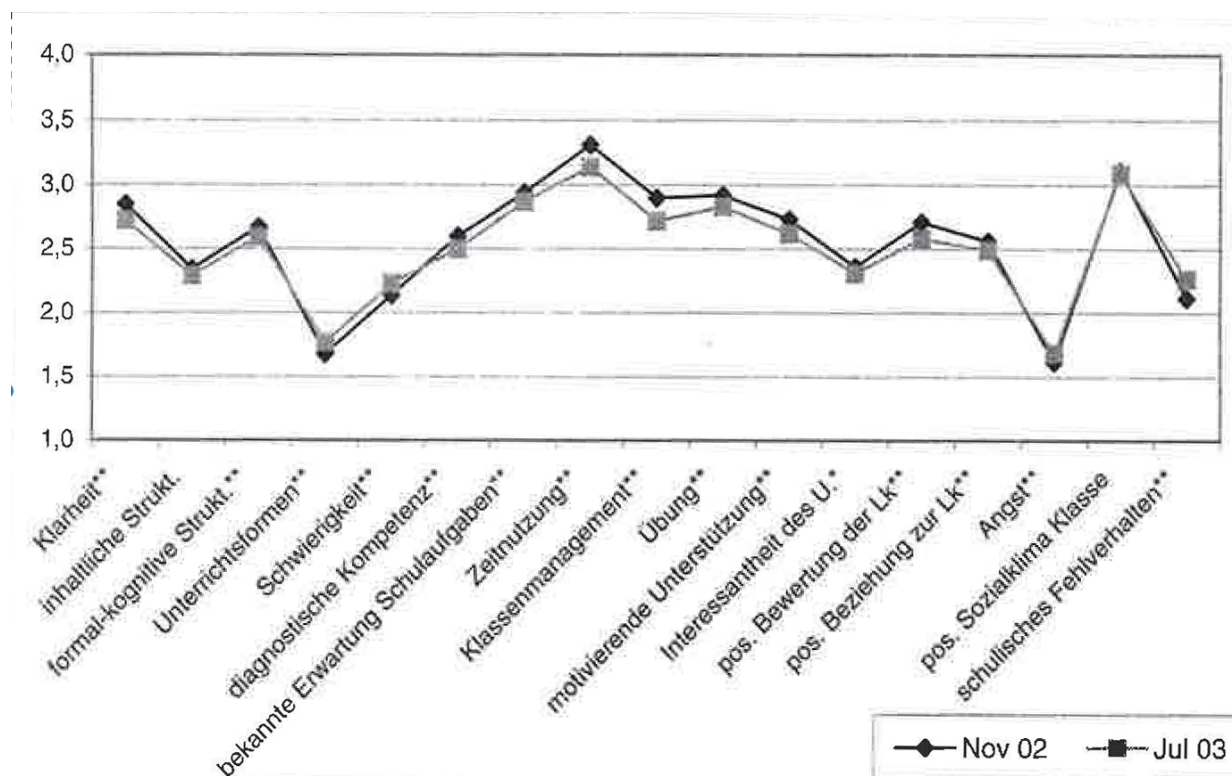


Abbildung 1: Prä-/Postvergleich Skalen der Unterrichtsqualität, Untersuchung Ditton und Arnold (2004), \* $p < 0,05$ ; \*\* $p < 0,01$

Beim Vergleich der Schülerwahrnehmungen der beiden Erhebungszeitpunkte, fiel zunächst auf, dass sich die Beurteilungen generell als sehr stabil erwiesen, ein weiteres Indiz für die Reliabilität von Schülerurteilen (siehe Abb. 1). Allerdings wurde der Unterricht zum zweiten Zeitpunkt in fast allen Skalen - zwar nur gering aber dennoch signifikant - schlechter beurteilt. Wie in Abschnitt 2.1.7 bereits beschrieben, führt Feedback alleine nicht automatisch zu einer Verbesserung von Unterricht, da sehr viele Begleitfaktoren den Regelkreis beeinflussen. Dies wurde nach einer genaueren Analyse der Daten noch weiter verdeutlicht, denn obwohl die Gesamtstichprobe relativ stabile Werte aufwies, zeigten sich bei lehrerspezifischer Betrachtung beträchtliche Unterschiede. Klassen, die von jungen Lehrern (unter 40 Jahre) unterrichtet wurden, Lehrpersonen, die eine hohe Selbstwirksamkeit besaßen, und solche, die mit der Ergebnisdarstellung (Mittelwertangaben, Darstellung von Interquartilsvergleichen) gut zu Recht kamen, zeigten im Gegensatz zur Gesamtstichprobe zum zweiten Messzeitpunkt eine Verbesserung der Beurteilung durch die Schülerinnen und Schüler.

Des Weiteren wurde festgestellt, dass Lehrpersonen vor allem dann die Schülerinnen und Schüler als kompetent und die Ergebnisse als verwendbar einschätzten, wenn diese selbst von den Schülerinnen und Schüler positiv beurteilt wurden und sie eine hohe berufliche Zufriedenheit mit geringer Belastungswahrnehmung aufwiesen (Ditton & Arnold 2004).

Kritisch anzumerken wäre hierbei, dass die beiden Erhebungszeitpunkte mindestens sieben Monate auseinanderlagen und es ist anzuzweifeln, dass die Ergebnisse in einem Kausalzusammenhang zum Feedbackverfahren stehen. Außerdem gaben einige Lehrpersonen der Untersuchung auch an, dass sie weitere Unterstützung im Umgang mit den Ergebnissen gebraucht hätten und mit der Darstellung der Beurteilungen schlecht zu Recht gekommen seien.

Im Gegensatz zu Ditton & Arnold (2004) führten Gärtner & Vogt (2013) keine Längsschnittuntersuchung durch, sondern erforschten ausschließlich das Verhalten der Lehrpersonen nachdem sie eine Schülerbefragung durchgeführt hatten. Diese Schülerbefragung fand im Rahmen einer freiwilligen Selbstevaluation statt, die durch das Selbstevaluationsportal (SEP) der Bundesländer Berlin und Brandenburg angeregt wurde. Aufgrund der Freiwilligkeit am SEP und an der Untersuchung teilzunehmen, kann diese Untersuchung nicht als repräsentativ für die Gesamtpopulation von Lehrkräften angesehen werden. Trotzdem ist diese Studie für die weitere Erforschung aufschlussreich, da sie Erkenntnisse darüber liefert, wie Lehrpersonen Schülerfeedback einschätzen, wie sie mit den Rückmeldungen umgehen und welche Probleme dabei auftauchen können.

Um das Verhalten der Lehrpersonen nach dem Erhalt der Rückmeldungen (wofür ebenfalls der Fragebogen nach dem QuAIT-Modell eingesetzt wurde) zu untersuchen, wurden Interviews durchgeführt, die sich auf das Rahmenmodell nach Helmke & Hosenfeld (2005) stützten. Dieses Modell enthält vier Phasen:

1. Rezeption: Nach Hosenfeld und Groß Ophoff (2007) müssen Lehrkräfte eine Rückmeldung zum Unterricht zunächst wahrnehmen und verstehen.
2. Reflexion: In einem zweiten Schritt können sie die Ergebnisse reflektieren, also nach Erklärungen für die vorgefundenen Ergebnisse suchen.
3. Aktion: Entsprechend der jeweiligen Erklärung werden nun spezifische Maßnahmen zur Optimierung durchgeführt.
4. Evaluation: Das Modell endet mit einer erneuten Evaluation der eingeleiteten Maßnahmen, wobei das Ergebnis dieser Evaluation wiederum als Rückmeldung und somit Startpunkt eines neuen Evaluationszyklus angesehen wird. (Gärtner 2013, S. 112)

In den Interviews gaben die Lehrpersonen an, dass sie bei entsprechender Kritik ihre Methodik umstellten, verschärft auf spezifische Verhaltensweisen achteten und das methodisch-didaktische Vorgehen transparenter machten. Des Weiteren empfanden sie positive Schülerurteile als Bestärkung für ihre Arbeit und die Mehrheit der Lehrkräfte nahm eine Verbesserung des Schüler-Lehrpersonen-Verhältnisses wahr.

Von allen befragten Lehrpersonen besprachen 89 % die Ergebnisse der Schülerbefragung in der Klasse und 60% im Kreise des Kollegiums. Üblicherweise wurden der Klasse alle Ergebnisse präsentiert und kritische Punkte, zum Beispiel Abweichungen von der Selbstwahrnehmung, besprochen. Dabei haben die Lehrpersonen die Reflexion des Feedbacks im Klassengespräch für weitere zielgerichtete Aktivitäten in der Unterrichtsentwicklung als sehr wichtig empfunden. Die meisten der befragten Lehrkräfte wiesen in ihrer Selbsteinschätzung eine hohe Übereinstimmung mit den Schülerurteilen auf. Problematisch an der hohen Übereinstimmung war allerdings, dass die Lehrpersonen in Bereichen, in denen sie ihre eigene Wahrnehmung bestätigt sahen, keine Verhaltensänderung initiierten, selbst wenn die Einschätzungen auf niedrigem Niveau lagen. Es wurde aber auch berichtet, dass die eingesetzten Items zu einer Reflexion des Unterrichts führten, unabhängig von den Ergebnissen der Rückmeldung. Die Lehrpersonen sahen das Feedback als eine positive Bestätigung ihres beruflichen Handelns (was jedoch aufgrund der Freiwilligkeit auf eine Positivauswahl der Probanden zurückgeführt werden kann) und als einen Beitrag zur Verbesserung des Schüler-Lehrpersonen-Verhältnis. Zu Veränderungen in den Unterrichtshandlungen kam es laut der Studie vor allem dann, wenn die Rückmeldungen neue Informationen lieferten (überraschend waren), die Übereinstimmung mit der eigenen Wahrnehmung eher gering waren und eine konstruktive Auswertung mit der Klasse und/oder dem Kollegium durchgeführt wurde.

Die Ergebnisse der bisherigen Untersuchungen legen also nahe, Schüler-Feedback nicht als isoliertes Verfahren einzusetzen. Eine Vorstellung und Diskussion der Ergebnisse in der Klasse wird sowohl auf Forscher- als auch auf Lehrpersonenseite als wichtig empfunden. Untersuchungen aus dem Hochschulbereich fordern weitere Maßnahmen wie Coaching, Weiterbildungen oder Beratungen, um stärkere Effekte durch Feedback zu erzielen. Allerdings war es im Rahmen dieser Untersuchung nicht möglich ein Coaching-Programm zu installieren, zumal ein

gezieltes Programm erst nach Kenntnis über die Mechanismen von Schülerfeedback sinnvoll integriert werden kann.

Die beiden beschriebenen Studien im Regelschulbereich geben Hinweise darauf, dass Lehrpersonen auf die Feedbacks mit Änderungen in der Unterrichtsentwicklung reagieren und ihre Unterrichtsmethodik anpassen. Ob diese Maßnahmen aber die Unterrichtsqualität verbessern, ist nicht abschließend geklärt, da über einen Zeitraum eines dreiviertel Jahres eine Reihe anderer Faktoren die Unterrichtsgestaltung und die Unterrichtsqualität beeinflussen können.

Es stellt sich somit die Frage, welche Wirkung sich auf die Entwicklung der Qualität des (Physik-) Unterrichts erzielen lässt, wenn ein Schülerfeedback innerhalb eines kürzeren Zeitraums mehrmals durchgeführt wird. Dabei wäre außerdem interessant, ob sich Hinweise bezüglich eines Zusammenhangs zwischen den auftretenden Effekten (z. B. Verbesserung der Unterrichtsqualität) und Einstellungen, Reaktionen und Verhaltensweisen z. B. im Rahmen der Unterrichtsentwicklung auf Lehrpersonenseite finden lassen. Zu diesem Zweck werden Teile der Forschungsansätze und -instrumente der beiden vorgestellten Arbeiten übernommen, in einem Prä-/Postdesign kombiniert (siehe Kap. 4) und durch ein neu entwickeltes Feedbackverfahren, das im Unterrichtsalltag in relativ zeitnahen Abständen eingesetzt werden kann, erweitert.

### **3 Entwicklung eines Schüler-Kurzfeedbacks als Mess- und Interventionsinstrument**

Möchte man ein Schüler-Feedbackverfahren integrieren, das alltagstauglich und universell in kurzen Abständen regelmäßig durchgeführt werden kann, steht man vor der Herausforderung, einerseits genügend aufschlussreiche Daten über den Unterricht zu generieren und andererseits den Fragebogen in seinem Umfang gering zu halten, sodass dieser am Ende einer Unterrichtsstunde in ca. fünf Minuten ausgefüllt werden kann.

Grundlegende Idee für die Gestaltung des Fragebogens war, den Lehrpersonen die Möglichkeit zu geben, *in die Köpfe* der Schülerinnen und Schüler zu schauen, also zu erfahren, ob ihr Physikunterricht Denk- und Lernprozesse in Gang gesetzt hat. Diese Herangehensweise birgt mehrere Vorteile: (1.) wie schon in Abschnitt 2.1.2.1 beschrieben, könnten Verzerrungen in der Wahrnehmung abgemildert werden; (2.) es wird ein Beitrag zu Hatties Forderung (2013) geleistet „Lernen sichtbar zu machen“ und (3.) könnte durch diese Art der Evaluation der Bewertungscharakter abgeschwächt werden und der Lehrperson fiel es leichter, Schülerfeedback nicht als „Beurteilungs-, sondern als Entwicklungsinstrument“ (Bastian 2007, S.15) zu sehen (die Schülerinnen und Schüler beschreiben ihr eigenes Verhalten, nicht das der Lehrperson).

Als theoretische Bezugspunkte für die Entwicklung des Fragebogens dienen ausgewählte Elemente des konstruktivistischen Lehr-Lern-Verständnisses, sowie das Modell der drei Basisdimensionen der Unterrichtsqualität, das im Rahmen der COACTIV-Studie (Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schülerinnen und Schülern) entwickelt wurde (vgl. Kunter & Voss 2011) und auf einen Vorschlag im Kontext der TIMS-Studie zurückgeht (vgl. Klieme, Schümer & Knoll 2001).

#### **3.1 Lernen als Konstruktionsprozess**

##### **3.1.1 Vom radikal-konstruktivistischen zum sozial-konstruktivistischen Lernverständnis**

Im radikal-konstruktivistischen Lernverständnis (vgl. von Glaserfeld 1992, beruhend auf der genetischen Erkenntnistheorie von Piaget 1970) geht man davon aus, dass Wissen und Gelerntes immer subjektiv ist, dass kein Mensch ein objektives Abbild der Wirklichkeit in sich trägt, sondern dass dieses Wissen durch personenbezogene Aspekte (Vorwissen, Vorerfahrungen, Einstellungen usw.) geprägt ist und durch den Kontakt mit der Umwelt selbst konstruiert wird. Lernen ist also ein Prozess, den nur das Individuum selbst durchführt. Die neuen Umwelterfahrungen werden dabei mit den bisherigen kognitiven Konzepten verglichen und eingeordnet. Falls

es zu Widersprüchen zwischen Umwelt und den subjektiven Konzepten kommt, wird versucht, die Konzepte so anzupassen, dass diese mit den Vorerfahrungen kongruieren.

Dieser Vorgang wird mit Piagets Prozessen der Assimilation (einpassen und erweitern der wahrgenommen Umwelt in bereits bestehende kognitive Strukturen) und Akkomodation (Veränderung der kognitiven Strukturen aufgrund von Widersprüchen mit der wahrgenommen Umwelt) in Verbindung gebracht. Nach dieser Vorstellung gibt es den Prozess des Lehrens und Vermittelns nicht, woraus sich wiederum die Annahme ableiten ließe, dass die Aufgabe der Lehrperson ausschließlich darin besteht, Lern-Umwelten zu schaffen, die das selbständige Konstruieren von Wissen ermöglichen. Dem sei jedoch entgegenzuhalten, dass es wahrscheinlich „einer Illusion gleichkommt, anzunehmen, Kulturwissen, zu dessen Gewinnung die Menschheit Jahrhunderte gebraucht hat, ließe sich im Unterricht ohne massive Anleitung nachentdecken“ (Reusser 2006, S.159).

Unter dem Einfluss u.a. von Vygotsky (1978), der im Gegensatz zu Piaget die Sprache nicht nur als Kommunikationswerkzeug des bereits Gedachten und Gelernten sondern als Werkzeug für das Denken und Lernen selbst sah, wurde der radikale Konstruktivismus um den sozialen (moderaten) Konstruktivismus erweitert: Soziale Kommunikation (auch das stumme Selbstgespräch) und die Lernsituation regen das Lernen an und beeinflussen es. Vygotsky beschreibt, dass durch soziale Interaktion die aktuelle kognitive Entwicklungszone (wie sie Piaget beschreibt) des Lernenden zur Zone der nächsten Entwicklung (Zone of proximal development, ZPD) erweitert werden kann.

Diese Art der sozialen Interaktion wurde im Laufe der Jahre mit dem Begriff des Scaffoldings in Verbindung gebracht (vgl. Wood, Bruner & Ross 1976) und charakterisiert sich dadurch, dass die Lehrperson u.a. durch das Vorgeben der Lernziele, der Kontrolle ihrer Einhaltung und der Komplexitäts-Reduktionen auf den wesentlichen Lerngegenstand, dem Lernenden ein Lerngerüst zur Verfügung stellt (scaffold engl. für Gerüst), das mit steigender Erfahrung und Fähigkeit des Lernenden schrittweise abgebaut wird.

Ein weiterer wichtiger Begriff im Kontext konstruktivistischer Lerntheorien ist der des *Conceptual Change*. Dieser stammt aus dem angelsächsischen Sprachraum und ist mit den Piaget'schen Begriffen Akkomodation und Assimilation verwandt (vgl. Möller 2007), konzentriert sich aber stärker auf die bereits vorhandenen Konzepte (Präkonzepte) von Lernenden bevor sie mit dem Lerngegenstand, dessen Erforschung oder der diesbezüglichen Lehrmeinung konfrontiert werden. Es hat sich herausgestellt, dass diese Präkonzepte sehr stabil sein können und die Lernenden daran hindern ihre Konzepte zu ändern oder zu erweitern (vgl. Duit 1996).

Zusammengefasst erkennt der soziale Konstruktivismus an, dass das Lernen ausschließlich vom Lernenden durchgeführt werden kann und sich in einem Erweitern und Verändern von bereits vorhandenen kognitiven Konzepten zeigt (Conceptual Growth bzw. Conceptual Change). Der soziale Konstruktivismus hebt sich jedoch vom radikalen Konstruktivismus ab, indem er

anerkennt, dass ein Lernprozess von sozialen Situationen beeinflusst werden kann und somit eine Steuerungsdimension eröffnet wird.

In der Interpretation des konstruktivistischen Lehr-/Lernprozesses kam es in der Vergangenheit häufig zu Diskussionen über dessen Auswirkung auf die angewandte Lehrmethodik: Wird der Konstruktivismus in seiner radikalen Form aufgefasst, ergeben sich daraus Lehrmethoden, die Kirschner, Sweller, & Clark (2006) unter dem Begriff des *minimally guided* zusammengefasst haben. Geht man aber von einem sozialen Konstruktivismus aus, kann dieser weitestgehend lehrmethodenunabhängig betrachtet werden (vgl. Hmelo-Silver et al. 2007). Denn das Konstruieren von Wissen bezieht sich nur auf den Prozess innerhalb der kognitiven Strukturen, dieser kann sowohl durch ein Gedankenexperiment, einen Vortrag oder das selbständige und forschende Handeln angeregt werden (vgl. Reusser 2006). Dies bedeutet jedoch nicht, dass die Auswahl der Methode irrelevant ist, sondern lediglich, dass die gängigen Lehr-Methoden bei adäquatem Einsatz ein *Conceptual-Change/Growth-Potential* in sich tragen. In Teilen der Literatur wird diesbezüglich auch zwischen Sicht- und Tiefenstrukturen unterschieden, also einerseits den methodischen Zugängen, Unterrichtsinszenierungen und Rahmenbedingungen und andererseits den Merkmalen des eigentlichen Lehr-/Lernprozesses (Kunter & Voss 2011). Die beiden Autoren zitieren Untersuchungen (u.a. Lipowsky 2002, Veenman, Kenter & Post 2000), die belegen, dass Sicht- und Tiefenstrukturen eines Unterrichts unabhängig voneinander auftreten können und dass die Tiefenstrukturen größere Erklärungsmacht bezüglich Lernzuwachs besitzen (Geller, Neumann & Fischer 2014, Hattie 2013, Seidel & Shavelson 2007, Wang, Haertel & Walberg 1993).

Den Einsatz von bestimmten Unterrichtsmethoden in einem kurz gehaltenen Fragebogen abzufragen, scheint deshalb nicht zielführend. Ganz im Gegensatz zur Erhebung von Einschätzungen zur kognitiven Aktivität und zu Lernprozessen.

### **3.1.2 Die Rolle von Experimenten beim Konstruieren neuen Wissens**

In diesem Zusammenhang nimmt das Experimentieren einen Sonderstatus ein, denn „das Experiment spielt im naturwissenschaftlichen Unterricht eine zentrale Rolle. Schließlich ist das Experiment Kern naturwissenschaftlicher Methoden zur Gewinnung von Erkenntnissen über die Natur“ (Duit, Tesch & Mikelskis-Seifert 2010, S. 1). Die *Fähigkeit der Erkenntnisgewinnung* wiederum, ist neben dem *Aufbau von Fachwissen*, der *Fähigkeit zur fachgemäßen Kommunikation* und der *kontextuellen Bewertung von naturwissenschaftlich geprägten Sachverhalten*, einer von vier Kompetenzbereichen der KMK-Bildungsstandards für den mittleren Bildungsabschluss (2004), beziehungsweise einer von sechs Handlungsaspekten der Schweizer Grundkompetenzen Naturwissenschaften (EDK 2011). Beachtenswert ist, dass das Experimentieren nicht nur die *Fähigkeit der Erkenntnisgewinnung* und den *Aufbau von Fachwissen* fördert, sondern auch

die *Fähigkeit zur fachgemäßen Kommunikation* und die *kontextuelle Bewertung von naturwissenschaftlich geprägten Sachverhalten* begünstigt, wenn sich Schülerinnen und Schüler über das experimentelle Vorgehen austauschen und in seinem entsprechenden Kontext bewerten.

Innerhalb der unterrichtsmethodischen Dimension sieht Labudde (2000, S. 119-120) in den Schüler- und Lehrerexperimenten vier Merkmale, die für eine gesonderte Stellung als konstruktivistisches Element sprechen:

1. *Das Zusammenspiel zwischen Handlung und Operation*: [...] [Es wird anerkannt], dass naturwissenschaftliche Begriffe und Theorien zwar wohl menschliche Konstruktionen sind, die aber in der ständigen experimentellen Auseinandersetzung mit den Objekten entwickelt, überprüft und allenfalls modifiziert werden [...].
2. *Intersubjektivität durch Experimente*: Experimente bieten für Lernende und Lehrende günstige Voraussetzungen, gemeinsame Erfahrungen zu machen und zu teilen. Die Jugendlichen stellen - zusammen mit ihren Lehrkräften, zu zweit oder in Gruppen - Hypothesen auf, beobachten, werten aus und interpretieren. [Damit schafft man] eine günstige Basis für die Konstruktion von Wissen [...].
3. *Bedeutung der Experimente für den naturwissenschaftlichen Unterricht*: Experimente gelten als wesentliches Charakteristikum des naturwissenschaftlichen Unterrichts [...].
4. *Bedeutung der Experimente für Naturwissenschaftslehrkräfte*: Wie empirische Untersuchungen zeigen, gehören Experimente zu den wichtigsten Orientierungspunkten für Lehrkräfte bei der Planung und Durchführung des Unterrichts (z. B. Fischler 1989) [...].

Labudde hebt hervor, dass das Experiment aus moderat konstruktivistischer Sicht als Bindeglied zwischen Theorie und Realität zu sehen ist. Die, im Wissenschaftsbetrieb angewandte, theoriegeleitete Vorgehensweise während des Experimentierprozesses sollte demnach auch für Schülerinnen und Schüler gelten (Labudde 2000). Ein zentraler Aspekt soll dabei das „Zusammenführen der individuellen Präkonzepte mit tragfähigen naturwissenschaftlichen Ideen“ darstellen (Börlin 2012, S. 37). Einem Präkonzept wird demnach ein realer und erfahrbarer Sachverhalt gegenübergestellt. Dies kann eine besonders starke Anregung für einen Conceptual-Change-Prozess darstellen. Allerdings zeigt Börlin (2012), dass Lehrpersonen zwar während des Experimentierprozesses auf die adäquate Verwendung von Fachtermini achten, jedoch eher selten die Schülerinnen und Schüler dazu anregen, die Experimente mit Erklärungsmodellen in Verbindung zu bringen.

Um diese Prozesse anzuregen, z. B. durch Versuch und Irrtum, ist es nötig den Schülerinnen und Schülern während des Experimentierens gewisse Freiräume zu geben. Andererseits birgt das offene Experimentieren (Abwesenheit jeglicher Steuerungs- oder Scaffoldingstrategien) die Gefahr, den Bezug auf das dahinterstehende theoretische Konzept zu verlieren. „Beim Einbezug von Handlungsformen ist es deshalb entscheidend, die Funktion von Handlungen für den Aufbau- und Verstehensprozess im Blick zu behalten“ (Adamina & Möller 2010, S. 106).

Möchte man dies für ein Schüler-Feedback fruchtbar machen, folgt daraus, dass die Lehrperson Informationen erhalten sollte, inwieweit das Experimentieren die intendierten theoretischen Vorstellungen verdeutlicht hat, also ob ein *Theoriebezug der Experimente* für die Schülerinnen



und Schüler erkennbar war. Denn das Durchführen oder Beobachten von Experimenten kann durchaus mit viel Freude und Engagement auf Schülerseite einhergehen, dies garantiert aber noch keine Erweiterung des naturwissenschaftlichen Verständnisses.

Und es lässt auch keine Schlüsse darauf zu, ob das Experiment auf eine, für den Prozess der Erkenntnisgewinnung, nützliche Weise eingesetzt wurde. Achten die Schülerinnen und Schüler zum Beispiel auf die Kontrolle von Störvariablen oder auf korrekte Hypothesenbildung und deren Überprüfung? Im angestrebten Kurzfeedback sollten die Lehrpersonen deshalb auch Informationen darüber erhalten, ob die Schülerinnen und Schüler die *wissenschaftliche Arbeitsweise* während der Durchführung von Schüler- aber auch von Demonstrationsexperimenten reflektiert und berücksichtigt haben.

### 3.2 Der Kurzfeedbackbogen

Diese Überlegungen werden nun mit den, im Rahmen der TIMS-Studie vorgeschlagenen (Klime, Schümer & Knoll 2001) und später in der COACTIV-Studie entwickelten Basisdimensionen, in Verbindung gebracht. Diese stehen im engen Zusammenhang mit dem aus der Sicht der Unterrichtsqualitätsforschung beschriebenen *Angebot-Nutzen-Paradigma* (siehe Abschnitt 4.1.1), das Unterricht als ein Angebot sieht, welches von Schülerinnen und Schülern innerhalb der Rahmenbedingungen individuenabhängig genutzt wird.

Nimmt man die sozial-konstruktivistische Sichtweise des Lernens als Grundlage, sollten die Lehrpersonen Informationen darüber erhalten, ob durch ihren Unterricht eine Anregung für einen Wissenszuwachs beziehungsweise ein erweitertes Konzeptverständnis bei den Schülerinnen und Schülern stattgefunden hat und ob diese eine angemessene Unterstützung bei der Erschließung der Lerninhalte erfuhren.

Diese beiden Prozesse stellen für Kunter & Voss (2011) zwei der drei Basisdimensionen dar. Die erste umschreiben sie als *kognitive Aktivierung* und die zweite als *konstruktive Unterstützung*. Wobei die konstruktive Unterstützung sowohl das Scaffolding beinhaltet als auch die Qualität der Schüler-Lehrpersonen-Beziehung. Da die Qualität der Schüler-Lehrpersonen-Beziehung jedoch eine relativ persistente Zustandsbeschreibung darstellt und somit nicht schulstundenspezifisch ist, erscheint es nicht sinnvoll, diese in einem Kurzfeedback innerhalb eines geschlossenen Antwortformats abzufragen.

Hingegen lohnt es sich, den Aspekt des Scaffoldings zu berücksichtigen. Ein gutes Scaffolding kann sich (a) in einer gut strukturierten Aufbereitung des Lerninhaltes beziehungsweise des Unterrichtssettings und (b) in der Angemessenheit, sowohl bezüglich der Stoffauswahl als auch bezüglich der Unterstützung während der Bearbeitung des Selbigen, bemerkbar machen (vgl. Kunter & Voss 2011), kurz: Ist der Unterricht (zu a) gut *strukturiert* und (zu b) für die Schülerinnen und Schüler *kognitiv passend*?

Darüber hinaus führen die Autoren noch eine dritte Basisdimension auf, nämlich die der effizienten Klassenführung, sie stellt sozusagen die Rahmenbedingung dar, in der Wissen konstruiert werden kann. Zahlreiche Untersuchungen belegen, dass die Leistungen der Schülerinnen und Schüler umso höher ist, je störungsärmer der Unterricht ist und je mehr Lernzeit zur Verfügung steht (Junge, Arx & Labudde 2014, Baumert et al. 2004, Seidel & Shavelson 2007 und Wang et al. 1993). Somit kann eine Rückmeldung über die, auf Schülerseite wahrgenommene Disziplin während des Unterrichts eine wichtige Information darstellen.

Fasst man diese Überlegungen und jene aus Abschnitt 3.1.2 zusammen, ergeben sich die folgenden Skalen, die sowohl die Basisdimensionen des Unterrichts abdecken als auch die spezifischen Besonderheiten des Experimentierens berücksichtigen:

- Kognitive Aktivierung
- Kognitive Passung
- Strukturiertheit
- Disziplin
- Theoriebezug der Experimente
- Bezug des wissenschaftlichen Arbeitens bei Experimenten

Zu diesen Skalen wurden für diese Studie Items entwickelt, welche die Selbsteinschätzung der Schülerinnen und Schüler auf einer vierstufigen Likert-Skala (4 = trifft voll zu, 1 = trifft gar nicht zu) erfassen sollen. Diese Items wurden in einer Vorstudie (drei Schulklassen, 74 Schülerinnen und Schüler) pilotiert und einer Faktorenanalyse und Reliabilitätsprüfung unterzogen. Auf dieser Grundlage wurden einzelne Items angepasst oder ergänzt. Daraus resultierte der folgende Kurzfeedback-Fragebogen für die Hauptstudie (siehe auch Anhang 8.1.1):

- *Kognitive Aktivierung (Cronbachs  $\alpha = 0,74$ )*
  - Der Unterricht hat mich dazu angeregt, genauer über das Unterrichtsthema nachzudenken.
  - Der Unterricht hat dazu geführt, dass ich mein bisheriges Wissen über das heutige Thema korrigiert oder erweitert habe.
  - Ich habe heute in Physik etwas dazugelernt.
  - Ich habe mich heute intensiv mit den Unterrichtsinhalten beschäftigt.
- *Kognitive Passung (nicht unterfordert) (Cronbachs  $\alpha = 0,75$ )*
  - Ich finde, dass man den Unterrichtsstoff schneller hätte durchnehmen können.
  - Ich fand den Unterrichtsinhalt zu leicht.
  - Heute wurden Dinge, die eigentlich klar waren, zu ausführlich besprochen.
  - Ich fand, dass es heute im Unterricht zu langsam vorwärts ging.

- *Kognitive Passung (nicht überfordert) (Cronbachs  $\alpha = 0,83$ )*
  - Ich bin heute im Unterricht nicht mitgekommen.
  - Ich fand den Unterrichtsinhalt zu schwer.
  - Ich hätte heute mehr Unterstützung gebraucht, um im Unterricht mitzukommen.
  - Mir fiel es heute schwer, dem Unterricht zu folgen.
- *Strukturiertheit (Cronbachs  $\alpha = 0,81$ )*
  - Für mich war klar, welche Inhalte im heutigen Unterricht gelernt werden sollen.
  - Für mich war klar, welches Lernziel heute erreicht werden sollte.
  - Ich bin heute Schritt für Schritt dem Lernziel näher gekommen.
  - Für mich ist klar, wie die Inhalte des heutigen Unterrichtsinhalts zusammenhängen.
  - Die Reihenfolge mit der die heutigen Inhalte durchgenommen wurden, verwirrte mich.
- *Disziplin (Cronbachs  $\alpha = 0,84$ )*
  - Um wirklich gut lernen zu können, war es heute im Unterricht zu laut.
  - Ich habe mir heute gewünscht, dass die Lehrperson früher für Ordnung gesorgt hätte.
  - Heute war im Unterricht eine gute Atmosphäre.
  - Die Lehrperson hatte die Klasse heute gut im Griff.
- *Theoriebezug der Experimente (Cronbachs  $\alpha = 0,84$ )*
  - Die heutigen Experimente haben mich dazu angeregt über den Unterrichtsinhalt genauer nachzudenken.
  - Die heutigen Experimente haben mir geholfen, die Herleitung der Theorie oder die Theorie selber, besser zu verstehen.
  - Die Experimente haben dazu geführt, dass ich mein bisheriges Wissen über das heutige Thema korrigiert oder erweitert habe.
  - Ich habe mir Gedanken gemacht, wie die Ergebnisse der Experimente zu erklären sind.
- *Bezug wissenschaftliches Arbeiten bei Experimenten (Cronbachs  $\alpha = 0,82$ )*
  - Ich habe während des Experimentierens darauf geachtet, ob genau gearbeitet wird.
  - Ich habe darauf geachtet, ob die Regeln zum richtigen Experimentieren eingehalten werden. (Bei diesem Item konnten die Schülerinnen und Schüler auch „kann ich nicht beantworten“ ankreuzen)
  - Während der Experimente habe ich mir überlegt, wie man vorgehen muss, um eine Hypothese (Vermutung) eindeutig zu überprüfen.

Außerdem wurde den Schülerinnen und Schülern die Möglichkeit gegeben in zwei offenen Antwortformaten Rückmeldung an die Lehrperson zu geben:

- Folgende Themen des heutigen Unterrichts sollten nochmals angesprochen werden, damit ich sie verstehe:
- Was ich unbedingt noch zu der heutigen Schulstunde sagen will:

Auf diese Weise konnten die Schülerinnen und Schüler erstens auf konkrete Schwierigkeiten bezüglich der Unterrichtsinhalte hinweisen und zweitens auf pädagogisch/didaktische Kritikpunkte aufmerksam machen (Angst, Motivation etc.), die nicht über die Skalen abgedeckt werden.

Es versteht sich von selbst, dass ein detaillierter Erkenntnisgewinn über die Güte des gehaltenen Unterrichts allein auf Grundlage des Fragebogens nicht möglich ist. Das Ziel liegt vor allem darin, erste Hinweise für den Erfolg des eigenen Unterrichts zu gewinnen, die über die Sichtstrukturebene hinausreichen, und damit eine Grundlage für tiefergehende Analysen während der Ergebnisvorstellung in der Klasse zu schaffen (siehe Abschnitt 4.4.1). Aus der Sicht des Angebot-Nutzen-Paradigmas (siehe oben und Abschnitt 4.2.1) soll das Schüler-Kurzfeedback Informationen liefern, in welchem Maße die Schülerinnen und Schüler das (Unterrichts-) *Angebot genutzt* haben. In einer möglichen Nachbesprechung sollen die Gründe für einen mangelnden *Nutzen* erörtert werden (z. B. war das *Angebot* nicht ausreichend oder die Rahmenbedingungen ungünstig?).

## 4 Die Studie

Die Studie verfolgt das Ziel, Hinweise bezüglich der Wirkung des oben vorgestellten Schüler-Kurzfeedbacks auf ausgewählte lehr-/lernprozessrelevante Aspekte zu finden und es auf seine Praxistauglichkeit zu prüfen, wenn dies in zeitlich nahen Abständen durchgeführt wird (Schüler-Kurzfeedback-Verfahren). Hierfür wird das Verfahren in ein Prä-/Postdesign eingebettet, die resultierenden Daten werden qualitativ und quantitativ ausgewertet, um abschließend empirisch gestützte Hypothesen hinsichtlich der Wirkung des Schüler-Kurzfeedback-Verfahrens auf ausgewählte Aspekte im Lehr-/Lernprozess zu formulieren.

Im Folgenden werden zunächst die Forschungsfragen formuliert, um daraufhin die zu untersuchenden, beeinflussbaren Faktoren und die dafür verwendeten Instrumente näher zu beschreiben. Im Anschluss wird das Forschungsdesign (Abb. 2) dargestellt. Abschließend erfolgt ein Ausblick auf die Auswertungsmethodik der ermittelten Daten.

### 4.1 Forschungsfragen

In den vorangegangenen Kapiteln wurde der theoretische Forschungsstand bezüglich Verarbeitung und Wirkung von Feedback im Allgemeinen und Schülerfeedback im Speziellen beschrieben, um daraufhin ein theoretisch fundiertes Interventions-/Erhebungs-Instrument zu entwickeln, das die Wirkung und Praxistauglichkeit von Schülerfeedback verbessern soll. Da Schülerfeedback, anders als zum Beispiel eine veränderte Unterrichtsmethode, auf einer Metaebene des Lehr-/Lernprozesses angreift, können auch jegliche Aspekte innerhalb dieses Prozesses beeinflusst werden. In dieser Studie wird eine Einschränkung vorgenommen, die sich, nach Einschätzung von hypothetisch erhöhter Beeinflussbarkeit durch Schülerfeedback, fachdidaktischer Relevanz und forschungsökonomischen Gesichtspunkten, herausgebildet hat. Folgende Aspekte sollen in der Studie Berücksichtigung finden:

- ausgewählte *Faktoren der Unterrichtsqualität* (aus Sicht der Schülerinnen und Schüler),
  - im Besonderen der Teilaspekt des *Schüler-Lehrpersonen-Verhältnisses* (aus Sicht der Schüler/-innen und Lehrpersonen);
- die *Unterrichtsentwicklung* und das *Unterrichtshandeln* der Lehrperson (aus Sicht der Schüler/-innen und Lehrpersonen);
- die Lehrpersonen-Einschätzung der *kognitiven und sozialen Voraussetzungen* (*Lernvoraussetzung, Leistungshomogenität der Klasse, Klassenmanagement*) und des *Beurteilungsvermögens von Unterricht* der Schülerinnen und Schüler.

Diese Aspekte werden nicht isoliert betrachtet, sondern es soll versucht werden, die gewonnenen Daten in einen plausiblen Wirkungszusammenhang zu bringen, also nach Mustern gesucht

werden, die innerhalb des Verfahrens positive Entwicklungen der lehr-/lernprozessrelevanten Faktoren begünstigen.

Dies führt zu folgenden vier Forschungsfragen (FF):

#### **Wirkung des Schüler-Kurzfeedback-Verfahrens**

- FF 1: Lassen sich Hinweise identifizieren bezüglich des Einflusses von Schüler-Kurzfeedback auf a) ausgewählte Skalen der Unterrichtsqualität unter besonderer Berücksichtigung des Schüler-Lehrpersonen-Verhältnisses und b) die Unterrichtsentwicklung und das Unterrichtshandeln der Lehrperson?
- FF 2: Welche Entwicklungen nehmen Schülerinnen und Schüler im Unterricht innerhalb des Schüler-Kurzfeedback-Verfahrens wahr?
- FF 3: Lassen sich Hinweise für Muster identifizieren, die eine positive Wirkung von Schüler-feedback auf die Unterrichtsqualität begünstigen?

#### **Praktikabilität des Schüler-Kurzfeedback-Verfahrens**

- FF 4: Wie schätzen Lehrpersonen und Schülerinnen und Schüler die Praktikabilität des Schüler-Kurzfeedback-Verfahrens ein?

### **4.2 Instrumente zur Erfassung beeinflussbarer Faktoren und ihre theoretische Fundierung**

Die in den Forschungsfragen enthaltenen abhängigen Faktoren werden in den nächsten Abschnitten theoretisch fundiert beschrieben und die verwendeten Instrumente vorgestellt. Handelt es sich dabei um neu entwickelte Instrumente, werden die Items in den jeweiligen Abschnitten aufgeführt, bei bereits bestehenden Instrumenten wird entsprechend auf den Anhang verwiesen.

## 10 Schulklassen (10 Lehrpersonen) im 9. Schuljahr Physik

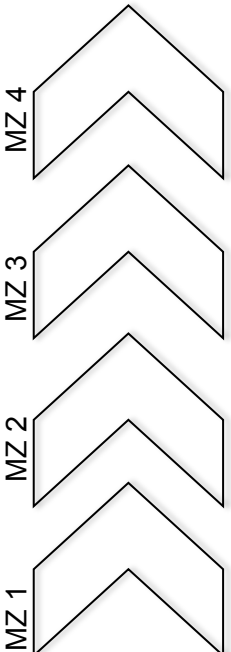
<p><i>t=0, Präerhebung</i></p> <p><b>Befragung der Schülerinnen und Schüler (SuS):</b></p> <ul style="list-style-type: none"> <li>• Unterrichtsqualität (angepasstes Instrument aus QuaSSU, Abschnitt 4.2.1)</li> </ul> <p><b>Befragung der Lehrpersonen (LP):</b></p> <ul style="list-style-type: none"> <li>• Selbsteinschätzung Unterrichtshandeln (Instrument aus SWISE-Studie, Abschnitt 4.2.2)</li> <li>• Beurteilung der sozial-kognitiven Voraussetzungen (Instrument neu entwickelt, Abschnitt 4.2.3)</li> <li>• Einschätzung des Beurteilungsvermögens von Unterricht der Schülerinnen und Schüler (Instrument neu entwickelt, Abschnitt 4.2.3)</li> <li>• LP-SuS-Verhältnis (Instrument neu entwickelt, Abschnitt 4.2.1.1)</li> </ul>	 <p><b>Intervention</b></p> <p>Alle 3-5 Unterrichtsstunden (4x) ein Kurzfeedback (neu entwickelt, siehe Unterkapitel 3.2) von Schülerinnen und Schülern an die Lehrperson. Diese gibt die Ergebnisse bekannt; weitere Thematisierung freiwillig.</p>	<p><i>t=10-12 Wochen, Posterhebung</i></p> <p><b>Befragung der Schülerinnen und Schüler:</b></p> <ul style="list-style-type: none"> <li>• siehe Präerhebung und</li> <li>• Wahrnehmung des Unterrichtshandelns der Lehrperson (Instrument neu entwickelt, Abschnitt 4.2.2)</li> <li>• Beurteilung des Feedbackverfahrens (Instrument neu entwickelt, Unterkapitel 4.3)</li> </ul> <p><b>Befragung der Lehrpersonen:</b></p> <ul style="list-style-type: none"> <li>• siehe Präerhebung und</li> <li>• Selbsteinschätzung des Unterrichtshandelns plus Angabe der Veränderung (Instrument aus SWISE-Studie, Abschnitt 4.2.2)</li> </ul> <p><b>Interview mit Lehrpersonen (Abschnitt 4.2.2 und 4.5.1):</b></p> <ul style="list-style-type: none"> <li>• Unterrichtsentwicklung/-handlung,</li> <li>• Schüler-Lehrpersonen-Verhältnis und</li> <li>• Beurteilung des Schüler-Kurzfeedback-Verfahrens</li> </ul>
--	--	--

Abbildung 2: Forschungsdesign

#### 4.2.1 Faktoren der Unterrichtsqualität

Der Aspekt *Unterrichtsqualität* stellt in dieser Studie nicht nur einen Teil des Rückmeldeprozesses während der Intervention dar, sondern er ist auch eine abhängige Variable in der Prä-/Posterhebung. Bereits in Unterkapitel 3.1 und Abschnitt 2.1.2 wurden bezüglich der *Unterrichtsqualität* die Perspektivenabhängigkeit, die zentralen Dimensionen und die relativ hohe Güte der Schülerbeurteilung beschrieben. Der letzte Punkt soll auch während des Prä-/Postvergleichs dienlich sein, um eine Veränderung in der *Unterrichtsqualität* im Zeitraum des Schüler-Kurzfeedback-Verfahrens beobachten zu können. Im Gegensatz zum Fragebogen des Kurzfeedbacks soll allerdings ein umfassenderes Bild von *Unterrichtsqualität* erschlossen werden.

In den vergangenen Jahrzehnten veränderten sich die Forschungsparadigmen bezüglich *Unterrichtsqualität* vom Persönlichkeits-Paradigma über das Prozess-Produkt-Paradigma hin zum Angebots-Nutzen-Paradigma (vgl. u.a. Gröschner & Kleinknecht 2013, Lotz 2016). Das Persönlichkeitsparadigma dominierte in den 1950er bis späten 1960er Jahren und stellte die Persönlichkeitseigenschaften der Lehrperson für den Lernerfolg bei Schülerinnen und Schülern in den Mittelpunkt.

Da jedoch die Ergebnisse diesbezüglich eher uneinheitlich ausfielen, konzentrierte man sich ab den 1960er Jahren stärker auf das Verhalten der Lehrpersonen während des Unterrichts (Gröschner & Kleinknecht 2013). Also darauf, welche Merkmale im Unterrichtsprozess zu finden sind, und wie diese mit guten Unterrichtsergebnissen (z. B. der Leistungszuwachs der Schülerinnen und Schüler innerhalb eines Schuljahres) zusammenhängen (Prozess-Produkt-Paradigma) (Gräsel & Göbel 2011). Obwohl dieser Ansatz die, in Unterkapitel 3.1 bereits erwähnten, Tiefenstrukturen vernachlässigt und die Sichtstrukturen oft additiv zusammenhangslos betrachtet, entstammen diesem Paradigma eine Fülle von Untersuchungen mit teils stabilen Ergebnissen (Ditton 2002).

Im Zuge der sozial-konstruktivistischen Sichtweise des Lehr-/Lernprozesses wurde das Prozess-Produkt-Paradigma zum Angebot-Nutzen-Paradigma erweitert. Hierbei wird einerseits die Art und Weise, wie die Lehrperson Lerngelegenheiten im Unterricht anbietet und andererseits, wie die Schülerinnen und Schüler dieses Angebot interpretieren und nutzen (also die Tiefenstrukturen), in die Betrachtung mit einbezogen (Gröschner und Kleinknecht 2013).

Aus diesem Paradigma sind zahlreiche empirisch gestützte Unterrichtsmodelle und Erhebungsinstrumente hervorgegangen, wie beispielsweise die bereits beschriebenen Basisdimensionen von Unterricht (siehe Unterkapitel 3.1) oder das Angebot-Nutzen-Modell nach Helmke (2009) und das daraus entwickelte Instrument EMU (Evidenzbasierte Methoden der Unterrichtsdiagnostik und -entwicklung) (Helmke et al. 2013). Dieses berücksichtigt unter anderem auch die Perspektivenabhängigkeit der Unterrichtsqualitätsbeurteilung (Helmke & Lenske 2013).



Das aus dem amerikanischen Sprachraum stammende Modell QuAIT (Slavin 1996) berücksichtigt ebenfalls die Tiefenstrukturen des Unterrichtsangebots und der jeweiligen Nutzung. Das Akronym QuAIT steht für:

1. "Quality of Instruction": Der Grad, zu dem Lehrinhalte und Informationen so präsentiert werden, dass Schüler sie leicht lernen können. Die Qualität des Unterrichts ist in der Hauptsache abhängig von der Qualität des Curriculums und der Präsentation bzw. Darstellung der Inhalte im Unterricht.
2. "Appropriateness": Der Grad, zu dem Lehrende sicherstellen, dass die Lernenden bereit sind, neuen Stoff zu lernen, d.h., dass notwendig vorauszusetzende Fähigkeiten und Wissen vorhanden sind, der Stoff aber nicht schon gelernt wurde. Angemessenheit ist somit dann gegeben, wenn der Unterricht weder zu schwer noch zu leicht für die Lernenden ist.
3. "Incentive(s)": Der Grad, zu dem Lehrende sicherstellen, dass die Lernenden motiviert sind dem Unterricht zu folgen, sich zu beteiligen und sich die Inhalte anzueignen.
4. "Time": Der Grad, zu dem die Lernenden ausreichend Zeit auf die Aneignung des Stoffes verwenden, was auf die zugestandene Zeit (allocated) und die effektiv genutzte Lernzeit (engaged time; time on task) verweist. Letzteres steht wiederum mit der Qualität des Unterrichts und der Motivation der Schüler in Verbindung. (Ditton 2000, S. 81-82)

Auf dieser Grundlage und unter zusätzlicher Berücksichtigung sozialer Faktoren entwickelten Ditton, Arnoldt & Bornemann (2002), innerhalb des Projekts QuaSSU (Qualitätssicherung an Schulen), einen Schülerfragebogen. Die Übersicht der Skalen sind in Abschnitt 2.2.2, Abb. 1 dargestellt. Um innerhalb dieser Arbeit eine Vergleichbarkeit mit den Studien von Ditton & Arnold (2004) und Gärtner & Vogt (2013) zu gewährleisten, wurden die Skalen dieses Fragebogens größtenteils übernommen.

Da für die Durchführung der Präbefragung und die Einführung der Schülerinnen und Schüler in das Feedbackverfahren insgesamt nur 45 min zur Verfügung standen, wurden Skalen entfernt, deren Änderungswahrscheinlichkeiten - aufgrund theoretischer Überlegungen - geringer angesehen wurden. Die in dieser Studie verwendeten Skalen und Items sind in Abschnitt 5.1.1, Tab. 3 und im Anhang 8.1.2 zu finden.

#### **4.2.1.1 Schüler-Lehrpersonen-Verhältnis**

Das *Schüler-Lehrpersonen-Verhältnis* aus Schülersicht stellt innerhalb des QuaSSU-Fragebogens eine Sub-Skala der *Unterrichtsqualität* dar. Aus zwei Gründen erscheint es interessant, dieser Skala eine besondere Aufmerksamkeit zu schenken.

Erstens eröffnet das Schülerfeedback einen weiteren Kommunikationsweg zwischen der Lehrperson und den Schülerinnen und Schülern, was einerseits zu einem besseren Verständnis der Handlungsweisen zwischen den beiden Parteien führen könnte; andererseits könnten sich durch die Kommunikation aber auch bereits bestehende Konflikte verschärfen.

Zweitens zeigen Metaanalysen (Hattie 2013 und Cornelius-White 2007), dass dieser Faktor für den Lernerfolg eine hohe Bedeutung aufweist. Hattie (2013) ermittelte in seiner Metaanalyse eine durchschnittliche Effektstärke von  $d = 0,72$ . Somit rangiert das *Schüler-Lehrpersonen-*

*Verhältnis* direkt hinter dem Faktor *Feedback* auf Rang elf der wirkmächtigsten Faktoren auf die Lernleistung.

Innerhalb dieser Studie werden sowohl die Schülerinnen und Schüler als auch die Lehrpersonen betreffend des Schüler-Lehrpersonen-Verhältnisses befragt (*Daten-Triangulation*, siehe auch Abschnitt 4.5.2). Die Items auf Lehrpersonenseite wurden neu entwickelt. Während der Hauptstudie wurde ein Item entfernt, um eine angemessene Reliabilität der Skala zu erreichen:

*Einschätzung Schüler/-innen-Lehrpersonenverhältnis* aus Lehrpersonensicht (vierstufige Likert-Skala, Cronbachs  $\alpha = 0,85$ ,  $n = 10$ )

- Das Lehrer-Schülerverhältnis ist gut
- Ich würde die Klasse vermissen.
- Ich bin froh wenn ich die Klasse nicht mehr unterrichten muss.

Außerdem wird in den Lehrpersonen-Interviews auf diesen Aspekt eingegangen (*Methoden-Triangulation*, siehe Unterkapitel 4.4 und Abschnitt 4.5.1).

#### **4.2.2 Unterrichtsentwicklung/Unterrichtshandeln der Lehrperson**

Die Lehrperson als *Regisseur* der Angebotsseite von Unterricht trägt einen sehr großen Anteil zum Unterrichtsgeschehen bei. Deshalb ist von Interesse, ob und in welchem Maße die Lehrperson in der Entwicklung des Unterrichts und im konkreten Unterrichtshandeln Änderungen aufgrund von Schülerrückmeldungen vornimmt, beziehungsweise, ob die Schülerinnen und Schüler mögliche Änderungen auf Nutzerseite wahrnehmen. Hierzu wurden sowohl die Schülerinnen und Schüler als auch die Lehrpersonen befragt (*Daten-Triangulation*, siehe Abschnitt 4.5.2).

Auf Schülerseite wurde zunächst ein einzelnes Item entwickelt, das nach dem ersten Teil der Hauptstudie durch die vier folgenden Items ersetzt wurde (siehe Abschnitt 5.2.4, Tab. 11):

- Die Lehrperson nahm das Schülerfeedback ernst. (vierstufige Likert-Skala: „stimme voll zu“ – „stimme gar nicht zu“)
- Sind dir seit dem Beginn der Feedbacks Veränderungen im Unterricht aufgefallen? (vierstufige Likert-Skala)
  - Falls du „stimme eher zu“ oder „stimme voll zu“ angekreuzt hast: Welche Punkte sind dir besonders aufgefallen? (offenes Format)
  - Glaubst du, dass die Veränderungen durch das Feedback ausgelöst wurden? (vierstufige Likert-Skala)

Auf Lehrpersonenseite kamen sowohl geschlossene Items innerhalb des Fragebogens bezüglich des sozial-konstruktivistisch orientierten *Unterrichtshandelns* als auch offene Fragen im

Interview bezüglich der *Unterrichtsentwicklung* und des *Unterrichtshandelns* zur Anwendung (*Methoden-Triangulation*, siehe Unterkapitel 4.4).

Von den 16 Items des Fragebogens (siehe Abschnitt 5.2.6, Tab. 20 und Anhang 8.1.4) zum konstruktivistisch orientierten Unterrichtshandeln wurden 13 aus dem Projekt SWISE (Swiss Science Education) (vgl. Koch, in Druck) entnommen. Koch (in Druck) entwickelte diese Items teilweise aus den Vorlagen von Rakoczy et al. (2005). Um ein noch etwas breiter angelegtes Bild des Unterrichtens zu erhalten, wurden drei weitere Items hinzugefügt:

In meinem Unterricht....

- finden ausführliche Klassengespräche statt.
- ist es mir sehr wichtig mit dem Stoff durchzukommen.
- achte ich besonders darauf, dass das richtige Fachvokabular benutzt wird.

Die in diesem Fall durchgeführte Selbstausskunft der Lehrperson ist der Kritik ausgesetzt, möglicherweise einen Bias der sozialen Erwünschtheit und der selbstdienlichen Verzerrung hervorzurufen (siehe Abschnitt 2.1.2.1 und 2.1.4). Dem steht allerdings gegenüber, dass es für den außenstehenden Beobachter sehr schwierig ist, die Tiefenstrukturen zu erkennen und ein Außenstehender oder Beobachter die Zielorientierung und Sinnhaftigkeit von Handlungen nur erschließen kann, „indem er das von ihm Beobachtete interpretiert“ (Mutzeck 2007, S. 693). Wo hingegen die handelnde Lehrperson die Motive und Absichten ihres Handelns selbst in sich trägt und somit auch am besten beschreiben kann (Koch, in Druck).

Die Postfragebögen der Lehrpersonen beinhalteten die Items zum *Unterrichtshandeln* in zweifacher Ausführung hinsichtlich der Ausprägungsbeschreibung. Die erste Ausführung war identisch zur Präerhebung, in der die Ausprägung der Items über eine vierstufigen Likert-Skala beschrieben werden konnten. Bei der zweiten Ausführung sollten die Lehrpersonen direkt angeben, ob das beschriebene *Unterrichtshandeln* in der Interventionsphase häufiger, gleich häufig oder seltener durchgeführt wurde (vgl. Abschnitt 5.2.6 und Anhang 8.1.5). Die Idee hinter dem Verfahren war, die gemessenen Differenzen aus der Prä-/Posterhebung mit den beschriebenen Handlungsveränderungen abzugleichen, um somit die Validität der Ergebnisse (insbesondere bei der Analyse einzelner Lehrpersonen) zu erhöhen. In der Forschungsliteratur wird dies als *Konvergenzmodell* der *Triangulation* beschrieben (Klieme & Bos 2000, vgl. Abschnitt 4.5.2). Im Laufe dieser Studie stellte sich heraus, dass die von den Lehrpersonen beschriebenen Veränderungen häufig im Widerspruch zu den Differenzen der Prä-/Posterhebung standen. Somit konnten die Daten nur sehr eingeschränkt für die Analyse des Lehrpersonenhandelns verwendet werden (vgl. Ergebnisse Abschnitt 5.2.6 und Diskussion Unterkapitel 6.1).

Neben der schriftlichen Befragung zum *Unterrichtshandeln* wurde auf die *Unterrichtsentwicklung* und das *Unterrichtshandeln* innerhalb der leitfadengestützten Interviews eingegangen. Um eine Vergleichbarkeit zur Studie von Gärtner und Vogt (2013) herzustellen, wurden die Interviews ebenfalls anhand des *Rahmenmodells der Unterrichtsentwicklung* nach Helmke und Ho-

senfeld (2005) geführt (siehe Abschnitt 2.2.2). Allerdings veränderte man den Fokus der ersten Phase (*Rezeption*). Die Lehrpersonen sollten nicht mehr danach befragt werden, welche Schüler-Rückmeldungen für sie zentral waren, sondern welche Affekte sie während der Rezeption bei sich selbst wahrnahmen. Dies geschah vor dem Hintergrund der Motivationsbildung, weil davon ausgegangen wird, dass nicht nur Kognitionen sondern auch Emotionen die Motivation beeinflussen (Hascher 2005). Und da die Lehrpersonen im Falle einer Schülerrückmeldung selbst zum Lernenden des eigenen Unterrichts werden, beeinflussen die Emotionen nicht nur die Motivation zum späteren Unterrichtshandeln sondern auch die Bereitschaft, sich in einem Lernprozess mit den Rückmeldungen zu beschäftigen.

„Grundsätzlich gilt es [jedoch] zu beachten, dass positive Gefühle wie Freude oder Stolz nicht zwingend lernfördernd wirken, ebenso wenig wie negative Gefühle wie Wut oder Trauer das Lernen zwingend blockieren müssen“ (Hascher 2005, S. 614). Sinclair (1988) weist z. B. nach, dass Menschen in schlechter Stimmung sich intensiver mit den Details von Informationen auseinandersetzen und dass diese spontane Kausalanalysen sozialer Situationen besser durchführen können (Abele 1985 und Bohner et al. 1988); somit unterlaufen Personen in schlechter Stimmung weniger fundamentale Attributionsfehler (Ross 1977).

Andererseits zeigen zahlreiche Untersuchungen eine erhöhte Kreativität, Effektivität und Flexibilität von Menschen in guter Stimmung (Bless & Fiedler 1999). So legten Isen, Daubmann & Nowicki (1987) dar, dass positiv gestimmte Personen bei Problemlöseaufgaben besser abschnitten und in einer anderen Studie von Isen & Means (1983) wurde festgestellt, dass diese in komplexen Entscheidungssituationen zwar nicht besser aber deutlich schneller zu Lösungen gelangten.

Abele (1999) beschreibt, dass „in negativem emotionalen Zustand die Motivation stark am Ziel der Emotionsbewältigung orientiert ist, die Aufgabenbearbeitung dient u.a. zu diesem Zweck. Die Motivation in positiver Stimmung ist variabler und hinsichtlich der Aufgabenbewältigung stärker intrinsisch orientiert“. Somit können sowohl positive als auch negative emotionale Zustände für eine stärkere Motivation hinsichtlich der Unterrichtsentwicklung führen. Es ist also von Interesse, ob sich Hinweise für ein Muster in der Emotion während der *Rezeption* auf das spätere Unterrichtshandeln und die Entwicklung der Unterrichtsqualität feststellen lassen. Dementsprechend wurden folgende Leitfragen für die Interviews formuliert:

*Rezeption:* Wie haben Sie die Ergebnisse des Feedbacks (emotional) aufgenommen?

*Reflexion:* Wie haben Sie die Ergebnisse der Feedbacks interpretiert/bewertet?

*Motivation:* Welche Motivationen haben sich aufgrund des Feedbacks entwickelt?/  
Welche Motivation ergab sich für zukünftiges Unterrichtshandeln nach der Intervention?

*Handlung:* Inwiefern haben Sie die Motivationen in Handlung umgesetzt?/ Welche Faktoren haben die Änderung der Unterrichtshandlung besonders begünstigt (Fragebogen/Diskussion)?

Diese Leitfragen wurden sowohl auf das Feedbackverfahren im Allgemeinen als auch im Hinblick auf auffällige Ergebnisse in den Kurzfeedbacks (*stimulated recall*) angewendet.

Zusätzlich wurden die Lehrpersonen gefragt, ob sie den Unterricht aufgrund des Wissens, dass ein Feedback folgt, in besonderer Form entwickelt haben.

#### **4.2.3 Lehrpersonen-Einschätzung der sozial-kognitiven Voraussetzungen und des Beurteilungsvermögens von Unterricht durch Schülerinnen und Schülern**

Ein Teilaspekt des Angebot-Nutzen-Modells von Helmke (2009) stellen die *Kontextmerkmale* der Klasse dar. Im Allgemeinen handelt es sich um Faktoren wie das oben bereits erwähnte *Schüler-Lehrpersonen-Verhältnis* und die eingeschätzten *sozial-kognitiven Voraussetzungen* innerhalb der Klasse (*Klassenmanagement*, *Lernvoraussetzungen* der Schülerinnen und Schüler und *Leistungs-Homogenität*). Die Lehrperson schätzt die Ausprägung dieser Faktoren ein und variiert aufgrund dieser ihr Unterrichtsangebot. Da ein Schülerfeedback die Kommunikation über Unterricht erweitert, wäre es möglich, dass die Lehrperson die *Kontextmerkmale*, die sie über eine Klasse besitzt, verändert oder, dass bereits entwickelte Einschätzungen den Umgang mit dem Schülerfeedback beeinflussen und dieses somit auf andere Faktoren (z. B. *Unterrichtsqualität*) mehr oder weniger wirksam wird.

Ein weiteres *Kontextmerkmal*, das im Falle eines Schülerfeedbackverfahrens eine wichtige Rolle spielen kann, ist die *eingeschätzte Unterrichtsbeurteilungsfähigkeit* der Schülerinnen und Schüler. Auch hier könnte sich sowohl die *Einschätzung der Unterrichtsbeurteilungsfähigkeit* verändern als auch ein Einfluss auf den Umgang mit dem Schüler-Feedback ergeben.

Um diesen möglichen Entwicklungen nachgehen zu können, wurden in der schriftlichen Prä-/Posterhebung Fragen aus dem QuaSSU-Itempool eingesetzt (*Lernvoraussetzungen* und *Leistungshomogenität*) und Items entwickelt, die sich auf die *Unterrichtsbeurteilungsfähigkeit* der Schülerinnen und Schüler und das *Klassenmanagement* bezogen (vierstufige Likert-Skala, s.o. und Anhang 8.1.2). Die Items der *Unterrichtsbeurteilungsfähigkeit* wurden in einer Vorstudie (n = 16) einer Faktorenanalyse unterzogen und auf interne Konsistenz überprüft. Bei der Skala *Selbsteinschätzung Klassenmanagement* geschah dies in der Hauptstudie (ein Item wurde entfernt):

*Eingeschätzte Unterrichtsbeurteilungsfähigkeit* (Cronbachs  $\alpha = 0,65$ , n = 16):

- Schülerinnen und Schüler sind, gemittelt auf Klassenebene, kompetente Beurteiler von Unterricht.

- Das Urteil von Schülerinnen und Schülern bezüglich des Unterrichts muss man ernst nehmen.
- Ich glaube, dass ein Schülerfeedback zu einer Verbesserung des Unterrichts führen kann.
- Schülerinnen und Schüler bewerten vor allem den Unterricht am besten, der für sie am bequemsten ist.
- Es besteht eine große Gefahr, dass Schülerinnen und Schüler das Feedback als Rach-einstrument benutzen.

*Selbsteinschätzung Klassenmanagement* (Cronbachs  $\alpha = 0,84$ ,  $n = 10$ ):

- In dieser Klasse ist es überdurchschnittlich laut.
- Diese Klasse ist relativ schwer zu disziplinieren.
- Ich habe die Klasse im Griff.
- Ich schaffe es kaum, mich in dieser Klasse durchzusetzen.

Während der Lehrpersonen-Interviews wurde ebenfalls auf den Aspekt der sozial-kognitiven Voraussetzungen eingegangen (siehe Abschnitt 5.2.7.2).

### **4.3 Instrumente zur Erfassung der Einschätzung des Schüler-Kurzfeedback-Verfahrens**

Ein weiteres zentrales Ziel dieser Forschungsarbeit besteht darin, ein Schüler-Kurzfeedback-Verfahren zu entwickeln, das arbeitsökonomisch im Unterrichtsalltag eingesetzt werden kann, auch im Hinblick auf die Möglichkeiten der mobilen Informationstechnologien, die einen aufwändigen Auswertungsprozess auf Seiten der Lehrperson obsolet machen.

Neben den arbeitsökonomischen Faktoren muss das Verfahren von den Lehrpersonen, aber auch von den Schülerinnen und Schülern (*personenbezogene Daten-Triangulation*, siehe Abschnitt 4.5.2), inhaltlich akzeptiert werden, damit es im Schulalltag eingesetzt wird.

Auf Seiten der Schülerinnen und Schüler wurden diesbezüglich drei geschlossene (vierstufige Likert-Skala, s.o.) und zwei offene Items entwickelt, die innerhalb des Post-Fragebogens zum Einsatz kamen:

- Ich fand es gut, der Lehrperson Feedback geben zu können.
- Die Anzahl der Feedbacks in den letzten Wochen war zu hoch.
- Ich hätte in den letzten Wochen gern noch häufiger Feedback gegeben.
- So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich gut, weil... (max. 3 Stichpunkte)
- So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich schlecht, weil... (max. 3 Stichpunkte)

Auf Seiten der Lehrpersonen wurden während des Interviews folgende Leitfragen gestellt:

- Wie beurteilen Sie im Allgemeinen das Verfahren des Schülerfeedbacks?
- Haben Sie festgestellt, dass die Schülerinnen und Schüler keine Lust mehr hatten, das Feedback auszufüllen? Woran machen Sie das fest? Wie haben Sie die Schülerinnen und Schüler motiviert?
- Welche Aspekte des Verfahrens würden Sie auf jeden Fall ändern?
- Welche Aspekte des Verfahrens würden Sie auf jeden Fall beibehalten?
- Angenommen, es gäbe ein einfaches elektronisches Verfahren für regelmäßiges Schülerfeedback. Würden Sie es einsetzen und in welcher Form (zeitlicher Abstand, Besprechung mit der Klasse)? Warum?/ Warum nicht?

Um ein detaillierteres Bild über den Umgang mit dem Feedbackverfahren zu erhalten, wurden die Lehrperson während des Interviews außerdem bezüglich der Faktoren, die innerhalb des Verfahrens positiv auf die Unterrichtsentwicklung wirken könnten, befragt:

- Beschreiben Sie, welche Faktoren innerhalb des Feedbackverfahrens eine Änderung der Unterrichtsentwicklung besonders begünstigt haben.
- Was war entscheidender für die Entwicklung, das Feedback selbst oder die Diskussion in der Klasse darüber? Warum/- nicht?
- Sind Sie der Meinung das Feedback würde auch ohne die Diskussion einen hohen Effekt erzielen? Warum/- nicht?
- Beziehungsweise bräuchte es gar kein Feedback, wenn man nur öfters mit der Klasse über den Unterricht diskutieren würde? Warum/- nicht?
- Inwiefern haben Sie den Unterricht anders entwickelt, weil Sie wussten, dass es ein Feedback geben würde?

#### **4.4 Forschungsdesign**

Wie aus den bisherigen Darstellungen abzuleiten ist, werden sowohl quantitative als auch qualitative Verfahren verknüpft, um die Wirkung von Schüler-Kurzfeedback und die Praxistauglichkeit des Erhebungsverfahrens explorativ zu untersuchen. Diese Vorgehensweise wird in der wissenschaftlichen Literatur als Mixed-Method-Ansatz bezeichnet (Gläser-Zikuda et al. 2012, Patton 2002, Tashakkori & Teddie 2003) und soll dazu dienen, die spezifischen Schwächen einzelner Methoden zu kompensieren (Gläser-Zikuda & Järvelä 2008, Gröschner, Seidel & Shalvelson, Hofmann et al. 2008), da im Schul- und Bildungsbereich verschiedene soziale Konstruktionsprozesse stattfinden, in denen keine einfachen Kausalitäten auftreten (Gläser-Zikuda

et al. 2012, Tippelt & Schmidt 2010). Mayring (2001) unterscheidet vier verschiedene Analyse-Möglichkeiten zur Integration qualitativer und quantitativer Methoden auf Designebene:

- Vorstudienmodell:* Eine qualitative Vorstudie wird zur Hypothesenbildung genutzt, die danach mit Hilfe einer quantitativen Studie getestet werden.
- Verallgemeinerungsmodell:* Große qualitative Studien (z. B. Feldforschungen, Fallanalysen) werden in gesammelter Form einer breiten Prüfung unterzogen.
- Vertiefungsmodell:* Quantitativ erhobene Daten werden qualitativ vertiefend analysiert (z. B. Fallanalysen), um deren Interpretierbarkeit zu verbessern.
- Triangulationsmodell:* Innerhalb des Mixed-Methods-Ansatzes wird die *Methoden-Triangulation* als eine Verschränkung der qualitativen und quantitativen Analyseschritte verstanden, die gegenseitig unterstützend wirken soll. Das Endergebnis stellt den Schnittpunkt der Einzelergebnisse dar. Neben der *Methoden-Triangulation* existieren weitere Triangulationsmöglichkeiten, von denen innerhalb dieser Arbeit, der *Daten-Triangulation* eine wichtige Rolle zukommt (siehe auch Abschnitt 4.5.2).

Diese Dissertation berührt drei dieser Modelle:

1. Die gesamte Arbeit lässt sich innerhalb des *Vorstudienmodells* als qualitative Vorstudie betrachten, deren Ergebnisse eventuell quantitativ in einer Nachfolgestudie untersucht werden.
2. Die Ergebnisse der quantitativ erhobenen Schüler-Kurzfeedbacks werden mit Hilfe eines *stimulated recall* vertiefend analysiert und
3. Die *Faktoren des Unterrichtshandelns*, des *Schüler-Lehrpersonen-Verhältnisses*, die *Einschätzung der sozial-kognitiven Voraussetzungen*, des *Beurteilungsvermögens von Unterricht* und die *Einschätzung des Schüler-Kurzfeedback-Verfahrens* werden in einer *Methoden-Triangulation* (Fragebogen und Interview) verschränkt.

Neben der Festlegung des Methodendesigns stellte sich hinsichtlich des Ablaufs der Intervention, also der Durchführung der Schüler-Kurzfeedbacks, die Frage, ob ein digitales einem analogen Verfahren vorzuziehen sei. Da die Zeitpunkte für die Durchführung der Schüler-Kurzfeedbacks nicht fest vorgegeben waren und teilweise in acht Klassen im selben Zeitfenster durchgeführt wurden, war es den Forschenden kaum möglich, bei den Erhebungen immer anwesend zu sein. Dieser Umstand hätte für ein digitales Verfahren gesprochen, zumal das anschließende Auswertungsprozedere erheblich weniger Aufwand generiert hätte. Allerdings wurde es als schwierig erachtet, möglichst vollständige Datensätze von den Schülerinnen und Schülern zu erhalten, weil entweder die entsprechende Hardware an den Schulen oder unter den Schülerinnen und Schülern nicht ausreichend vorhanden war, um zum Unterrichtsende die



Schüler-Kurzfeedbacks möglichst umfassend durchzuführen. Somit entschied man sich für ein analoges Schüler-Kurzfeedback-Verfahren.

#### 4.4.1 Ablauf der Haupt-Studie

Die Hauptstudie (Abb. 2, S. 39) umfasste insgesamt einen Zeitraum von ca. 10-12 Wochen. Innerhalb dieses Zeitfensters wurden die Präerhebung, vier Schülerkurzfeedbacks, vier Präsentationen der Ergebnisse der Schülerkurzfeedbacks und die Posterhebung inklusive der Interviews mit den Lehrpersonen durchgeführt. Die Hauptstudie fand in zwei unterschiedlichen Zeitfenstern statt. Einmal zu Schuljahresende zwischen Mai und Juli (zwei Schulklassen) und zu Schuljahresmitte zwischen November und Februar (acht Schulklassen). Der erste Teil der Hauptstudie war zunächst als Pilotstudie gedacht, da jedoch nur sehr wenig Änderungen (siehe Abschnitt 5.2.4) in der Datenerhebung vorgenommen wurden, entschied man sich, die Daten für die Hauptstudie zu nutzen. Die insgesamt 10 Schulklassen wurden von 10 verschiedenen Lehrpersonen im Fach Physik unterrichtet. Sie stammten je zur Hälfte aus deutschen Gymnasien und Realschulen (mittlerer Bildungsabschluss) und befanden sich im 9. Schuljahr.

Nachdem die Einverständniserklärung der Eltern eingeholt war (siehe Anhang 8.2), wurde während der Präerhebung, mit Hilfe eines Fragebogens, auf Seiten der Schülerinnen und Schüler die *Unterrichtsqualität* (inkl. des *Schüler-Lehrpersonen-Verhältnisses*), die *demographischen* und die *Schulleistungs-Daten* abgefragt.

Auf Seiten der Lehrpersonen wurden eine Selbsteinschätzung zum *Unterrichtshandeln*, die Einschätzung der *sozial-kognitiven Voraussetzungen* sowie der *Beurteilungsfähigkeit von Unterricht* der Schülerinnen und Schüler und das *Schüler-Lehrpersonen-Verhältnis* ermittelt.

Außerdem wurde den Schülerinnen und Schülern das komplette Schüler-Kurzfeedback-Verfahren dargelegt und die Diagramme, die bei der Ergebnisvorstellung innerhalb der Intervention zur Anwendung kamen, erklärt.

Mit Hilfe der Präerhebungsdaten wurden die Fragebögen der Schüler-Kurzfeedbacks personalisiert um gegebenenfalls Rückschlüsse auf einzelne Schülerinnen und Schüler während der Ergebnisauswertung ziehen zu können. Daraufhin sendete das Forscherteam die Schüler-Kurzfeedback-Bögen in fünffacher Ausführung (inkl. einem jeweiligen Ersatzbogen) und vier frankierte Rückumschläge mit handschriftlicher Kennzeichnung an die Lehrpersonen.

In der folgenden Stunde wurde gegen Unterrichtsende das Schüler-Kurzfeedback ausgefüllt, von ausgewählten Mitschülern eingesammelt, im gekennzeichneten Rückumschlag versiegelt (um die Anonymität zu wahren) und der Lehrperson übergeben. Diese wiederum sendete den Umschlag an die Forschungsgruppe, welche die Daten auswertete und die offenen Antworten, die Klassenmittelwerte, inklusive der Standardabweichung, wieder zurück an die Lehrperson übermittelte (siehe Anhang 8.3). Parallel erhielt die Lehrperson nochmals schriftliche Erläute-

rungen bezüglich der dargestellten Diagramme und statistischen Kennwerte. Somit wurde der Forderung der Lehrpersonen, mehr Unterstützung hinsichtlich der Ergebnisdarstellung zu erhalten (Ditton & Arnold 2004), Rechnung getragen.

Die Lehrpersonen waren angehalten, in der nächsten Schulstunde die Ergebnisse der Befragung vorzustellen und bei Bedarf diese in der Klasse zu diskutieren (vgl. Abschnitt 2.2.2 ). Erst in der darauffolgenden (Doppel-)Stunde sollte dann das nächste Schülerfeedback stattfinden.

Nach vier Feedbacks und Auswertungs-Besprechungen wurde die Posterhebung durchgeführt. Die Schülerinnen-und-Schüler-Fragebögen beinhalteten erneut Fragen über die ausgewählten Skalen der *Unterrichtsqualität* sowie über das beobachtete *Unterrichtshandeln* der Lehrperson und über deren *Einschätzungen* bezüglich des *Schüler-Kurzfeedback-Verfahrens*.

Neben der schriftlichen Befragung fand zum Zeitpunkt der Posterhebung das leitfadengestützte Interview mit der Lehrperson statt, dessen Fokus auf der *Unterrichtsentwicklung/-handlung*, der *sozial-kognitiven Voraussetzungen*, dem *Schüler-Lehrpersonen-Verhältnis* und der *Beurteilung des Schüler-Kurzfeedback-Verfahrens* lag (vgl. Abschnitt 4.2.2, 4.3 und 4.5.1).

## **4.5 Auswertungsmethodik**

Nach der Haupterhebung lagen Daten aus zwei Perspektiven (Schüler/-innen und Lehrpersonen) in drei verschiedenen Formen vor:

1. Antworten geschlossener Items (Schülerinnen und Schüler und Lehrpersonen)
2. Antworten offener Items (Schülerinnen und Schüler)
3. Interviewdaten (Lehrpersonen)

Die Antworten der geschlossenen Items wurden mit Hilfe der Statistiksoftware SPSS hinsichtlich der Forschungsfragen ausgewertet, die jeweiligen Verfahren werden in den diesbezüglichen Ergebnisunterkapiteln (Kapitel 5) eingehender erläutert. Die offenen Antwortformate wurden in qualitativer und quantitativer Form analysiert, indem aus dem Material induktiv Kategorien gebildet und daraufhin ihre Häufigkeiten gezählt wurden. Jeweilige nähere Erläuterungen finden sich in den Abschnitten 5.2.4.2 und 5.4.2.

Der dritte Punkt wird im folgenden Abschnitt dargestellt. Daraufhin wird die methodische Verknüpfung der gewonnenen Daten erläutert (*Daten-Triangulation*).

### **4.5.1 Qualitative Inhaltsanalyse der Interviews**

Die Auswertung der Interviews wurde nach der Methode der strukturierenden qualitativen Inhaltsanalyse von Mayring (2010) durchgeführt. Diese besteht aus insgesamt acht Schritten (Abb. 3):

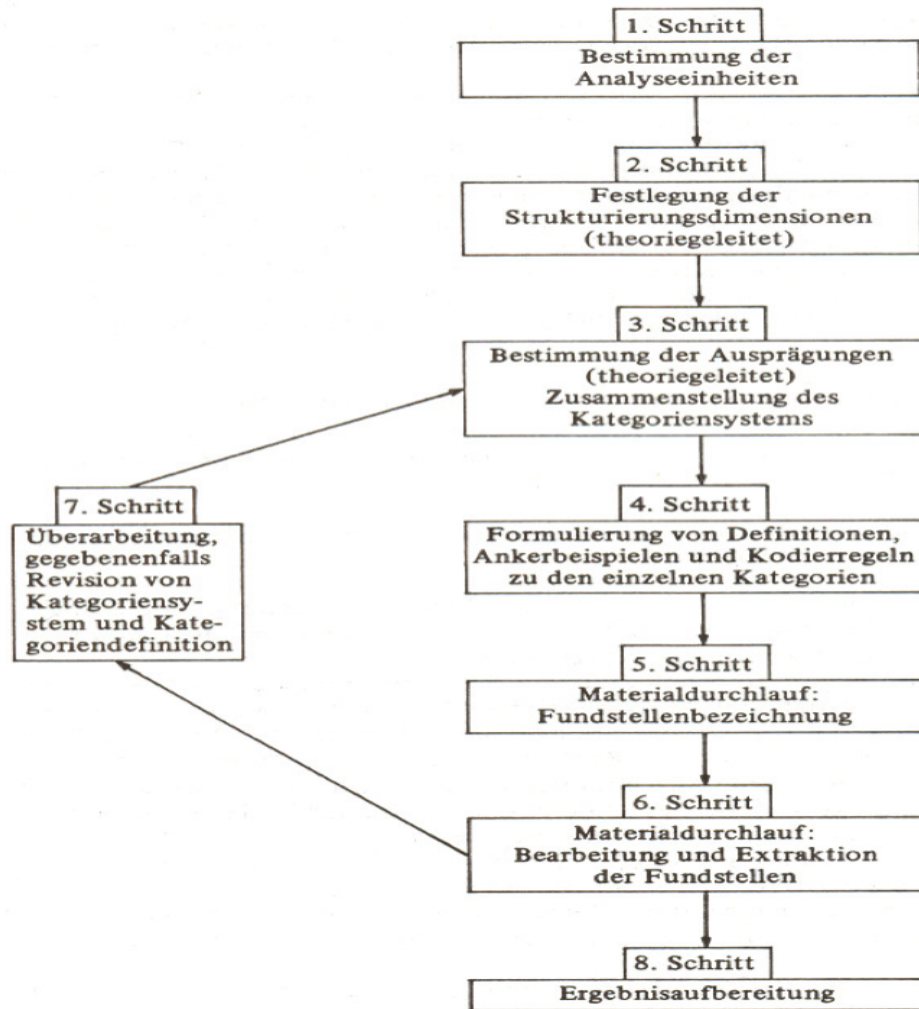


Abbildung 3: Ablaufmodell strukturierender Inhaltsanalyse (Mayring 2010, S. 93)

Dabei kann die Strukturierung unterschiedliche Ziele verfolgen. Innerhalb dieser Arbeit werden drei der vier von Mayring (2010) dargestellten möglichen Ziele verfolgt: die inhaltliche (Zusammenfassen von Themen und Inhalten aus dem Material), die typisierende (markante Ausprägungen im Material) und die skalierende (Ausprägungen von einzelnen Dimensionen im Material) Strukturierung. Hierzu wurde ein theoriegeleitetes Kategoriensystem erstellt, das im Laufe des Auswertungsprozesses angepasst wurde (deduktives und induktives Vorgehen, Tabelle 2):

Tabelle 2: Auswertungskategoriensystem der Interviews

<b>Unterrichtsentwicklung allgemein</b>	
<b>Rezeption</b>	emotional positiv
	emotional negativ
	nüchtern
<b>Interpretation</b>	Schüleraussage richtig aber nicht veränderbar
	Schüleraussage richtig und veränderbar
	Schüleraussage falsch
<b>Motivation</b>	Änderungsabsicht vorhanden
	keine Änderungsabsicht
<b>Handlung</b>	Handlung verändert
	Handlung unverändert
<b>Motivation für zukünftiges Unterrichtshandeln</b>	Beschreibung der Handlungsabsicht
<b>Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen, <i>stimulated recall</i>)</b>	
<b>Rezeption</b>	emotional positiv
	emotional negativ
	nüchtern
<b>Interpretation</b>	Schüleraussage richtig aber nicht veränderbar
	Schüleraussage richtig und veränderbar
	Schüleraussage falsch
<b>Motivation</b>	Änderungsabsicht vorhanden
	keine Änderungsabsicht
<b>Handlung</b>	Handlung verändert
	Handlung unverändert
<b>Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?</b>	ja
	nein
<b>Wahrgenommene Reaktionen der SuS auf das Feedback</b>	
<b>Störungsverhalten der SuS</b>	verbessert
	gleich geblieben
	verschlechtert
<b>Einschätzung der sozialen Voraussetzungen und der Beurteilungsfähigkeit von Unterricht</b>	
<b>Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht</b>	gestiegen
	gleich geblieben
	gesunken
<b>Allgemeine Einschätzung der SuS bzgl. der Reife</b>	gestiegen
	gleich geblieben
	gesunken
<b>Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS</b>	gestiegen
	gleich geblieben
	gesunken
<b>Allgemeine Einschätzung des SuS-LP-Verhältnisses</b>	verbessert
	gleich geblieben
	verschlechtert
<b>Einschätzung des Verfahrens Schülerkurzfeedback</b>	
<b>Das Verfahren insgesamt</b>	gut
	problematisch
	schlecht
<b>Generelle Einschätzung der Fragebögen</b>	gut
	schlecht
<b>Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse</b>	gut
	schlecht
<b>In elektronischer Form sinnvoll</b>	ja
	nein
<b>Würde es selbst in elektronischer Form einsetzen</b>	ja
	nein
<b>Rhythmus in dem die LP es einsetzen würde</b>	Intervall
<b>Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren</b>	

Alle zehn Interviews wurden von zwei Forschern codiert (einer der Codierer war der Verfasser des Kategoriensystems), verglichen und im Diskurs eine Konsensversion erstellt, die jedoch nur geringfügig von den zwei zuvor erstellten Codierungen abwich. Zwei Punkte führten während der Vergleiche häufiger zu Diskursen, die folgendermaßen gelöst wurden:

- Falls die Lehrperson Handlungen beschrieb, ohne explizit die vorangegangene Motivation zu nennen, wurde die, für die Handlung notwendige, vorausgehende Motivation entsprechend ergänzt.
- Beschrieb eine Lehrperson innerhalb des Rezeptionsprozesses eine Emotion mit dem Adjektiv „überraschend“ wurde im Diskurs geklärt, ob dieser Begriff im Sinne der Emotionsbeschreibung positiv, negativ oder neutral auszulegen ist.

Die Codierungen der einzelnen Interviews finden sich im Anhang 8.4.

Die zum Teil dichotomen Kategorien werden im Laufe der Auswertung expliziter und differenzierter dargestellt, um die Typisierungen und die Wirkungs-Muster innerhalb der Gruppen zu schärfen.

#### 4.5.2 Triangulation der Daten

Methoden-Triangulation als ein Verfahren innerhalb des Mixed-Methods-Ansatzes wurde bereits in Unterkapitel 4.4 angesprochen. Triangulation ist jedoch nicht ausschließlich als ein Aspekt des Mixed-Methods-Ansatzes zu sehen, auch wenn dies häufig formuliert wird (Kuckartz 2014). Flick (2000, S. 310) definiert Triangulation vereinfacht als „die Betrachtung eines Forschungsgegenstands von (mindestens) zwei Punkten aus“. Diese Betrachtung ist somit nicht auf die Dimension der verschiedenen Methoden beschränkt, sondern kann auch innerhalb anderer Dimensionen stattfinden. Denzin (1970) unterscheidet hierbei die *Methoden-*, *Daten-*, *Forscher-* und *Theorien-Triangulation*. Neben der bereits beschriebenen *Methoden-Triangulation* spielt die *Daten-Triangulation* innerhalb dieser Arbeit eine wichtige Rolle. Hierbei werden zur Untersuchung des selben Phänomens mannigfache Datenquellen einbezogen, welche sich wiederum in Zeit, Raum oder Person unterscheiden können (Denzin 1970).

Im vorliegenden Falle liegt eine *personenbezogene Daten-Triangulation* vor, da Daten hinsichtlich der Entwicklung des *Schüler-Lehrpersonen-Verhältnisses*, der *Praktikabilität des Schüler-Kurzfeedbackverfahrens* und seiner *Wirkung* sowohl auf Lehrpersonen als auch auf Schülerinnen- und Schüler-Seite erhoben wurden. Dabei wird das Ziel der *Komplementarität* verfolgt, also eine Vertiefung und Erweiterung der Erkenntnisse bezüglich der beschriebenen Forschungsgegenstände (Fielding & Fielding 1986). Triangulation kann aber auch dazu benutzt werden, um *Konvergenz* herzustellen, also die Erhöhung der Validität von Forschungsergebnissen aufgrund der verschiedenen Sichtweisen auf den Forschungsgegenstand.

Die Verfahren der unterschiedlichen Triangulationsmethoden, die in dieser Dissertation zur Anwendung kommen, sollen hauptsächlich dazu dienen, durch eine differenzierte Beschreibung von Mustern und Gruppen die Hypothesengenerierung zu präzisieren und empirisch zu untermauern. Hieraus können auch Handlungsempfehlungen für eine mögliche quantitative Nachfolgestudie entwickelt werden.

## 5 Ergebnisse

Die folgende Ergebnisauswertung soll in fünf Schritten erfolgen: Zunächst (1) wird die Reliabilität der Messinstrumente überprüft und eventuell Anpassungen beschrieben (z. B. Ausschluss von Items). Daraufhin (2) folgt eine deskriptive Übersicht der Daten über die gesamte Stichprobe zu den Messzeitpunkten Prä und Post. Im nächsten Schritt (3) kommt es zu einem klassenweisen Vergleich der Unterrichtsqualität als Grundlage für eine Gruppenbildung, die mit Hilfe weiterer Daten aus der Hauptstudie (4) untermauert wird. Zum Schluss (5) werden in einem explorativen Ansatz alle qualitativen und quantitativen Daten mit dem Ziel der Musteridentifizierung innerhalb der einzelnen Gruppen kombiniert, um daraus im folgenden Diskussionskapitel Hypothesen zu möglichen Wirkmechanismen von regelmäßigem Schülerfeedback ableiten zu können.

### 5.1 Statistische Güte der Skalen

Nachdem die Daten erhoben wurden, überprüfte man die Skalen (erneut) auf ihre innere Konsistenz. Im Folgenden werden diesbezüglich relevante Abweichungen von den vorgängigen Studien (Unterrichtsqualitäts-Skalen aus Schülersicht) sowie weitere Aggregationen von einzelnen Skalen zu Meta-Skalen (Schüler-Kurzfeedback) dargestellt.

#### 5.1.1 Reliabilität der Skalen zur Unterrichtsqualität aus Schülersicht

Die ausgewählten Skalen zur Unterrichtsqualität wiesen zu beiden Messzeitpunkten größtenteils ähnliche Werte für Cronbachs  $\alpha$  auf wie in der Originalliteratur (QuaSSU-Projekt) beschrieben ([www.quassu.net](http://www.quassu.net), abgerufen am 05.12.2016). Ausnahmen bildeten die Skalen „Angemessenheit Schwierigkeit“, „Zeitnutzung“ und „Inhaltliche Strukturiertheit“. Durch Entfernung je eines Items („Die Hausaufgaben sind so leicht, dass ich gar nicht wirklich nachdenken muss.“ und „Wie häufig kommt es vor, dass der Unterricht ausfällt?“) konnte für die beiden erstgenannten Skalen ein zufriedenstellendes Cronbachs  $\alpha \geq 0,65$  erreicht werden. Die Skala „Inhaltliche Strukturiertheit“ erreichte zu Prä einen Cronbachs  $\alpha$ -Wert von 0,55, zu Post einen Wert von 0,66. Da der Wert zum Post zufriedenstellend ist, wurde die Skala für die weiterführenden Analysen verwendet. Bei der Interpretation dieser Skala ist aber eine gewisse Vorsicht geboten. Tabelle 3 zeigt eine Übersicht der Reliabilitätskennwerte der angepassten Skalen.

Tabelle 3: Reliabilitätsstatistik der ausgewählten und angepassten Skalen zur Unterrichtsqualität aus Schülersicht

	QuaSSU	Prä	Post	
Skala	Cronbachs- $\alpha$	Cronbachs- $\alpha$	Cronbachs- $\alpha$	Anzahl der Items
Verhältnis zur Lehrperson	0,82	0,76	0,82	6
Klarheit der Lehrperson	0,76	0,65	0,80	3
Inhaltliche Strukturiertheit	0,69	0,55	0,66	4
Formal kognitive Strukturiertheit	0,72	0,68	0,77	5
Angemessene Schwierigkeit	0,78	0,77	0,73	4
Angemessenes Tempo	0,65	0,74	0,77	3
Diagnostische Kompetenz	0,90	0,86	0,88	12
Klassenmanagement	0,78	0,86	0,84	5
Bewertung der Lehrperson	0,91	0,83	0,89	6
Zeitnutzung	0,73	0,66	0,70	7
Interessantheit	0,85	0,78	0,82	9
Motivierende Unterstützung	0,74	0,65	0,76	4

### 5.1.2 Reliabilität der Skalen des Schüler-Kurzfeedbacks

Die gemäß Erkenntnissen aus der Vorstudie weiterentwickelten Skalen des Schülerfeedbacks (siehe Kap 3.2) ergaben gute Cronbachs- $\alpha$ -Werte (siehe Tab. 4).

Tabelle 4: Reliabilitätsstatistik Skalen der Schülerkurzfeedbacks

Skala	Cronbachs- $\alpha$	Anzahl der Items
Kognitive Passung "nicht zu leicht"	0,75	4
Kognitive Passung "nicht zu schwer"	0,83	4
Kognitive Aktivierung	0,74	4
Struktur	0,81	5
Disziplin	0,84	5
Theoriebezug Experimente	0,84	4
Bezug wissenschaftliches Arbeiten	0,82	3

Es wurde in der Folge mittels Faktorenanalyse überprüft, ob eine zusätzliche inhaltliche Aggregation der Items möglich ist. Die grundsätzliche Überlegung bestand darin, die Rezeption der Schülerinnen und Schüler bezüglich kognitiver und struktureller Elemente des Unterrichts im Allgemeinen (die Items der ersten vier Skalen der Tab. 4) und der Experimente im Speziellen (die Items der letzten beiden Skalen der Tab. 4) zusammenzufassen. Hierzu wurden zwei Faktorenanalysen durchgeführt (siehe Tabellen 6 und 7). Wie Tabelle 5 zeigt, luden die Items, die ein allgemeindidaktische Handeln beschreiben sollten, nicht auf einen gemeinsamen Faktor, sondern auf drei verschiedene. So konnten lediglich die Items der Skalen *kognitive Aktivierung* und *Struktur* zur Skala *Kognitiv aktivierende und strukturierende Elemente* zusammengefasst werden, wobei hier drei von fünf Items der Skala *Struktur* (mit \* markiert) auf einen zweiten Faktor luden. Allerdings sind die Faktorladungen auf den jeweilig anderen Faktor leicht schwächer, was für die oben durchgeführte Skalenbildung spricht.



Tabelle 5: Faktorenanalyse zur Dimensionsreduktion in allgemeindidaktisches Handeln

Rotierte Komponentenmatrix <sup>a</sup>			
	Komponente		
	1	2	3
Der Unterricht hat mich dazu angeregt, genauer über das Unterrichtsthema nachzudenken.		0,737	
*Für mich war klar, welche Inhalte im heutigen Unterricht gelernt werden sollen.	0,407	0,409	-,333
Der Unterricht hat dazu geführt, dass ich mein bisheriges Wissen über das heutige Thema korrigiert oder erweitert habe.		0,759	
Für mich war klar, welches Lernziel heute erreicht werden sollte	0,277	0,584	-,274
Ich habe heute in Physik etwas dazugelernt.		0,734	
Ich bin heute Schritt für Schritt dem Lernziel näher gekommen.	0,340	0,705	
Ich habe mich heute intensiv mit den Unterrichtsinhalten beschäftigt.	0,251	0,592	
*Für mich ist klar, wie die Inhalte des heutigen Unterrichts zusammenhängen.	0,531	0,555	
*Die Reihenfolge, mit der die heutigen Inhalte durchgenommen wurden, verwirrte mich.		0,276	0,213
Ich bin heute im Unterricht nicht mitgekommen.	0,670		
Ich finde, dass man den Unterrichtsstoff schneller hätte durchnehmen können.			0,778
Ich fand den Unterrichtsinhalt zu schwer.	0,809		
Ich fand den Unterrichtsinhalt zu leicht.	-,309		0,680
Ich hätte heute mehr Unterstützung gebraucht, um im Unterricht mitzukommen.	0,722	0,269	
Heute wurden Dinge, die eigentlich klar waren, zu ausführlich besprochen.			0,709
Mir fiel es heute schwer, dem Unterricht zu folgen.	0,821	0,244	
Ich fand, dass es heute im Unterricht zu langsam vorwärts ging.			0,811

Extraktionsmethode: Analyse der Hauptkomponente.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Rotation konvergierte in 5 Iterationen. \* = Lädt auf zwei Faktoren ähnlich hoch, Ladungen &lt; 0,2 ignoriert

Einfacher gestaltete sich die Ermittlung eines gemeinsamen Faktors der kognitiv methodischen Elemente des Experimentierens (siehe Tab. 6). Hier luden alle Items auf einen gemeinsamen Faktor, der als Skala *Experimentieren als theoriegeleitetes Handeln* definiert wurde.

Tabelle 6: Faktorenanalyse zur Dimensionsreduktion in Experimentieren als theoriegeleitetes Handeln

Komponentenmatrix <sup>a</sup>	
	Komponente 1
Die heutigen Experimente haben mich dazu angeregt über den Unterrichtsinhalt genauer nachzudenken.	,735
Ich habe während des Experimentierens darauf geachtet, ob genau gearbeitet wird.	,770
Die heutigen Experimente haben mir geholfen, die Herleitung der Theorie oder die Theorie selber, besser zu verstehen.	,704
Ich habe darauf geachtet, ob die Regeln zum richtigen Experimentieren, eingehalten werden.	,786
Die Experimente haben dazu geführt, dass ich mein bisheriges Wissen über das heutige Thema korrigiert oder erweitert habe.	,780
Ich habe mir überlegt, wie man vorgehen muss, um eine Hypothese (Vermutung) eindeutig zu überprüfen.	,805
Ich habe mir Gedanken gemacht, wie die Ergebnisse der Experimente zu erklären sind.	,740

Extraktionsmethode: Analyse der Hauptkomponente.

a. 1 Komponenten extrahiert.

Die Skala *Disziplin* wurde unverändert übernommen. Die Cronbachs  $\alpha$ -Werte der neuen, aggregierten Skalen sind als gut einzustufen (siehe Tab. 7). In allen weiteren Analysen wird mit diesen Skalen gearbeitet

Tabelle 7: Reliabilitätsstatistik der aggregierten Skalen des Schülerkurzfeedbacks

Skala	Cronbach-ALehrpersonha	Anzahl der Items
Kognitive Passung "nicht zu leicht"	,75	4
Kognitive Passung "nicht zu schwer"	,83	4
Kognitiv aktivierende und strukturierende Elemente	,86	9
Disziplin	,84	5
Experimentieren als theoriegeleitetes Handeln	,88	7

## 5.2 Ergebnisse bezogen auf die gesamte Stichprobe

Dieses Unterkapitel soll zunächst einen Überblick über die demographische Zusammensetzung der Stichprobe sowohl auf Schülerinnen und Schüler- und Lehrpersonenseite geben, um daraufhin, bezogen auf die Forschungsfragen FF 1, FF 2 und FF 4, die Ergebnisse der Prä-/Posterhebung und der Schüler-Kurzfeedbacks zu beschreiben. Zum Schluss erfolgt eine überblickgebende Auswertung der leitfadengestützten Interviews.

### 5.2.1 Demographische Daten der Schülerinnen und Schüler

Tabelle 8 zeigt die deskriptive Übersicht der demographischen Daten der Schülerinnen und Schüler. Diese stammen aus jeweils fünf Realschul- und fünf Gymnasialklassen des 9. Schuljahres. Die Klassenstärken lagen zwischen 19 und 28 Schülerinnen und Schülern, dabei ist der Anteil der Schülerinnen gegenüber den Schülern leicht erhöht. Etwa 80 % gaben Deutsch als Muttersprache an. Im arithmetischen Mittel erhielten die Schülerinnen und Schüler im Zeugnis des Vorjahres die Note 2,8. Da an Baden-Württembergischen Realschulen zu diesem Zeitpunkt Physik integriert in dem Fach Naturwissenschaftliches Arbeiten (NWA) unterrichtet wurde, handelt es sich nur bei den Schülerinnen und Schülern der Gymnasialklassen explizit um eine Physiknote. Gaben die Schülerinnen und Schüler im Fragebogen Viertelnoten an, so wurden diese für folgende Darstellung auf halbe Noten aufgerundet. Das Durchschnittsalter der Schülerinnen und Schüler betrug zum Zeitpunkt der Präerhebung 15,6 Jahre.

Tabelle 8: Demographische Daten der Schülerinnen und Schüler

Geschlecht				
		Häufigkeit	Prozent	Gültige Prozent
Gültig	weiblich	137	57,3	58,1
	männlich	99	41,4	41,9
	Gesamtsumme	236	98,7	100,0
Fehlend	kein Kreuz	3	1,3	
Gesamtsumme		239	100,0	
Ist Deutsch Deine Muttersprache?				
		Häufigkeit	Prozent	Gültige Prozent
Gültig	deutsch	189	79,1	80,8
	Fremdsprache	45	18,8	19,2
	Gesamtsumme	234	97,9	100,0
Fehlend	Doppelkreuz	2	0,8	
	kein Kreuz	3	1,3	
	Gesamtsumme	5	2,1	
Gesamtsumme		239	100,0	
Welche Note hattest Du im letzten Zeugnis in Physik?				
		Häufigkeit	Prozent	Gültige Prozent
Gültig	1,00	10	4,2	4,7
	2,00	63	26,3	29,9
	2,50	7	2,9	3,3
	3,00	85	35,5	40,2
	3,50	4	1,6	1,8
	4,00	37	15,4	17,5
	4,50	3	1,3	1,4
	5,00	2	0,8	0,9
	Gesamtsumme	211	88,3	100,0
Fehlend	kein Kreuz	28	11,7	
Gesamtsumme		239	100,0	
Alter				
		Häufigkeit	Prozent	Gültige Prozent
Gültig	13	3	1,3	1,3
	14	127	53,1	54,0
	15	76	31,8	32,3
	16	27	11,3	11,5
	17	1	0,4	0,4
	19	1	0,4	0,4
	Gesamtsumme	235	98,3	100,0
Fehlend		4	1,7	
Gesamtsumme		239	100,0	

### 5.2.2 Demographische Daten der Lehrpersonen

An der Studie nahmen zehn Lehrpersonen teil (neun männliche und eine weibliche). Das Altersspektrum lag zwischen 30 und > 60 Jahren, wobei die Hälfte der Lehrpersonen zwischen 30 und 39 alt waren. Dementsprechend verhält sich auch die Tätigkeitszeit im Lehrberuf: Nur drei der zehn Lehrpersonen können auf mindestens zehn Jahre Berufserfahrung zurückblicken. Die meisten unterrichten mehr als 20 Stunden pro Woche, von denen maximal vier in der untersuchten Klasse abgehalten werden. Sechs Lehrpersonen unterrichten nur das untersuchte Fach in der Klasse und vier ein weiteres Fach. Vier Lehrpersonen haben die Klasse in dem Schuljahr übernommen, in dem die Untersuchung durchgeführt wurde, sechs kennen die Klasse länger als 1,5 Jahre, eine Lehrperson sogar länger als vier Jahre. Eine Lehrperson unterrichtet fachfremd (Klasse 5). Die Tabellen 9 und 10 fassen die demographischen Daten sowohl einzeln für die Klasse/Lehrperson, als auch in ihrer Gesamtheit zusammen. Da es sich um zehn untersuchte Klassen handelt, lassen sich die Zahlenverhältnisse in den jeweils letzten Zeilen der Tabellen leicht ablesen.

Tabelle 9: Demographische Daten der Lehrpersonen. \*Ma=Mathe, NWT=Naturwissenschaft und Technik, NP=Naturphänomene, ITG=Informationstechnologie, Tech.=Technik, Engl.=Englisch, Ge.=Geschichte, GK=Gemeinschaftskunde, EK=Erdkunde, WK=Wirtschaftskunde, BNT=Biologie, Naturphänomene, Technik, Ch.=Chemie, Ph.=Physik

Lehrperson der Klasse	Geschlecht		Alter					Berufserfahrung (Jahre)						Neben Physik unterrichtete Fächer*
	m	w	20-29	30-39	40-49	50-59	> 60	1-2	3-5	6-10	11-20	21-30	> 30	
1	x				x					x				Ma, ITG
2	x						x						x	Ma.
3	x				x							x		Ma., NWT, NP, ITG
4		x			x							x		Ma.
5	x			x					x					Ma., Tech.
6	x			x						x				Ma., Engl., ITG
7	x			x						x				Ma., NP
8	x			x					x					Ma., Ge.
9	x			x					x					Ma., NWA, Tech., GK, EK, ITG, WK
10	x				x					x				Bio., Ch., NWA, BNT
Anzahl gesamt	9	1	0	5	4	0	1	0	3	4	0	2	1	

Tabelle 10: Demographische Daten der Lehrpersonen (Fortsetzung). \*Ma=Mathe, NWT=Naturwissenschaft und Technik, NP=Naturphänomene, ITG=Informationstechnologie, Tech.=Technik, Engl.=Englisch, Ge.=Geschichte, GK=Gemeinschaftskunde, EK=Erdkunde, WK=Wirtschaftskunde, BNT=Biologie, Naturphänomene, Technik, Ch.=Chemie, Ph.=Physik

Lehrperson der Klasse	Pensum (Stunden)					Pensum in der Klasse		Unterrichtszeit in der Klasse (Jahre)						Studierte Fächer*
	1-5	6-13	14-19	20-26	>26	<3	3-4	0,5-1	1-1,5	1,5-2	2-3	3-4	>4	
1				x		x				x				Ma., Ph., ITG
2				x			x				x			Ma., Ph.
3				x			x	x						Ma., Ph.
4		x				x						x		Ph.
5					x	x				x				Ma., Tech.
6				x		x						x		Ma., Ph., E.
7				x			x	x						Ph., Ma.
8				x		x		x						Ma., G., Ph.
9					x	x		x						Ma., Ph., WK, ITG
10				x			x						x	Bio., Ch., Ph.
Anzahl gesamt	0	1	0	7	2	6	4	4	0	2	1	2	1	

### 5.2.3 Unterrichtsqualität aus Schülersicht

Eine zentrale Fragestellung betrifft die möglichen Wirkungen von Schülerfeedback auf ausgewählte Aspekte der Unterrichtsqualität. Für die Auswertung wurden die Schülerantworten zum QuaSSU-Fragebogen so kodiert, dass vier eine hohe und eins eine geringe Zustimmung bedeutet. Die Daten, erhoben zu Prä und Post, wurden mit dem Statistikprogramm SPSS einer Varianzanalyse mit Messwiederholung unterzogen und ein Post-Hoc Test zur Identifizierung von Signifikanzen durchgeführt (Bonferroni-Methode, siehe Tab. 11).

Alle Skalenmittelwerte liegen über dem mittleren Wert von 2,5. Die Schülerinnen und Schüler beurteilen also die abgebildeten Bereiche der Unterrichtsqualität im Durchschnitt als eher gut oder gut. Hierbei stechen vor allem die Skalen *Zeitnutzung*, *Bewertung der Lehrperson* und *Klarheit der Lehrperson* heraus, die zu beiden Messzeitpunkten Werte über 3,1 aufweisen. Dagegen weisen die Skalen *Inhaltliche Strukturiertheit* und *Interessantheit* zu beiden Messzeitpunkt Werte unter 2,9 auf. Betrachtet man die Standardabweichungen der einzelnen Skalen fallen die überdurchschnittlich hohen Werte bei den Skalen *Angemessenheit Tempo* und *Klassenmanagement* auf. Hier scheinen entweder die Wahrnehmungsunterschiede zwischen den Schülerinnen und Schüler oder die Interklassenunterschiede leicht höher zu sein als bei den anderen Skalen.

Tabelle 11: Mittelwertvergleich Unterrichtsqualität aus Schülersicht über alle Klassen aggregiert. AM= arithmetischer Mittelwert, SD= Standardabweichung, SE= Standardfehler, Sig. = Signifikanzniveau \*p ≤ 0,05; \*\*p ≤ 0,01

Skala	AM Prä	AM Post	SD Prä	SD Post	SE Prä	SE Post	Post - Prä	Sig. p	Effektstär- ke d
<b>Verhältnis zur Lehrperson*</b>	2,80	2,88	0,57	0,66	0,041	0,047	0,08	0,03	0,14
<b>Klarheit der Lehrperson</b>	3,35	3,38	0,50	0,63	0,036	0,044	0,04	0,38	0,07
<b>Inhaltliche Strukturiertheit**</b>	2,59	2,74	0,56	0,62	0,040	0,044	0,15	0,00	0,27
<b>Formal kognitive Strukturiertheit**</b>	2,89	3,00	0,55	0,61	0,039	0,043	0,10	0,01	0,19
<b>Angemessenheit Schwierigkeit</b>	3,03	3,00	0,56	0,60	0,040	0,042	-0,03	0,47	-0,05
<b>Angemessenheit Tempo**</b>	2,95	3,09	0,74	0,72	0,053	0,051	0,15	0,00	0,20
<b>Diagnostische Kompetenz**</b>	2,80	2,89	0,54	0,58	0,038	0,041	0,09	0,01	0,17
<b>Klassenmanagement</b>	3,12	3,08	0,80	0,79	0,056	0,056	-0,04	0,22	-0,05
<b>Bewertung der Lehrperson</b>	3,17	3,19	0,57	0,73	0,040	0,051	0,02	0,56	0,04
<b>Zeitnutzung</b>	3,24	3,24	0,44	0,45	0,031	0,032	0,01	0,74	0,02
<b>Interessantheit**</b>	2,75	2,86	0,52	0,56	0,037	0,039	0,11	0,00	0,21
<b>Motivierende Unterstützung durch Lehrperson**</b>	3,05	3,16	0,53	0,61	0,037	0,043	0,11	0,00	0,21

Ermittelt man die Mittelwertunterschiede zwischen Prä und Post so ist festzustellen, dass zehn von zwölf Skalen eine positive Tendenz (Post - Prä ist positiv, drittletzte Spalte) aufweisen (Tab. 11 und Abb. 4). Sieben Skalen (*Inhaltliche Strukturiertheit*, *Formal kognitive Strukturiertheit*, *Angemessenheit Tempo*, *Diagnostische Kompetenz*, *Interessantheit* und *Motivierende Unterstützung durch Lehrperson*, *Verhältnis zur Lehrperson*) weisen signifikante positive Unterschied zwischen Prä und Post auf (vorletzte Spalte). Signifikant negative Unterschiede zwischen Prä und Post sind, gemittelt über die gesamte Stichprobe, nicht festzustellen.

Aus der Differenz der Mittelwerte und der gepoolten Standardabweichung (mit  $s_{gepoolt} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$ ) von Prä und Post wurden die zugehörigen Effektstärken berechnet (mit  $d = \frac{AM_{Post} - AM_{Prä}}{s_{gepoolt}}$ ). Diese sind insgesamt – aggregiert über alle Klassen – gering. Lediglich vier Skalen weisen nach der Definition von Cohen (1988) einen geringen Effekt auf ( $d > 0.2$ ). Dass die Effektstärken zwischen den untersuchten Klassen stark variieren, sei hier vorweggenommen und wird in den Unterkapiteln 5.4 und 5.5 eingehender analysiert.

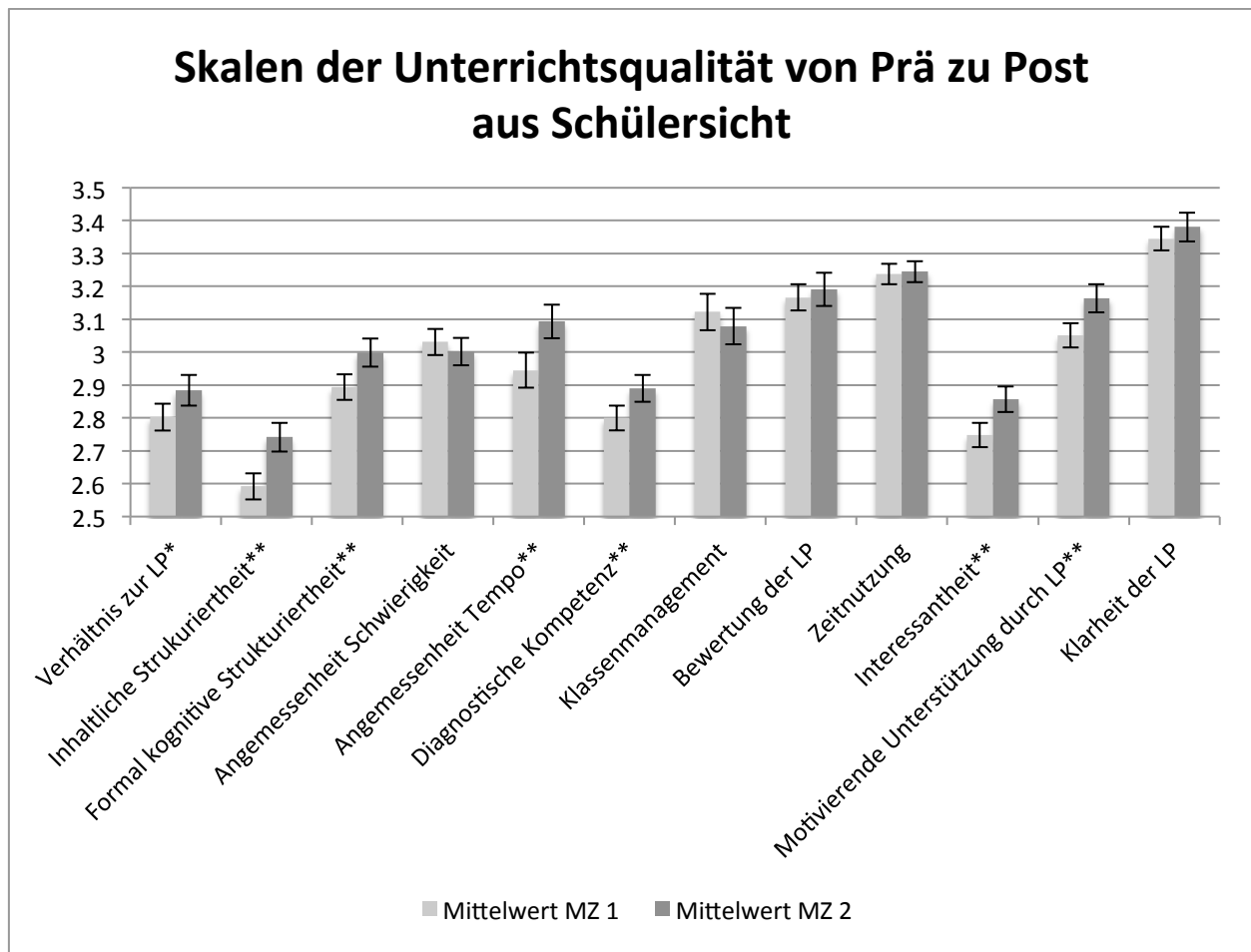


Abbildung 4: Mittelwertvergleich Unterrichtsqualität aus Schülersicht über alle Klassen aggregiert, Standardfehler als Fehlerbalken, Signifikanzniveau \* $p \leq 0.05$ ; \*\* $p \leq 0.01$

Insgesamt lässt sich also festhalten, dass über die gesamte Stichprobe hinweg ein klarer positiver Trend bei der Einschätzung einiger Skalen der Unterrichtsqualität durch die Lernenden zu beobachten ist und dass die ermittelten Effektstärken klein sind.

#### 5.2.4 Wahrnehmung des Schüler-Kurzfeedback-Verfahrens aus Schülersicht

Neben der Wahrnehmung der Unterrichtsqualität wurden die Schülerinnen und Schüler in der Posterhebung auch zu ihrer Einschätzung des Feedbackprozesses befragt. Neben einer Anzahl geschlossener Items mit 4-stufiger Likert-Skala (Kodierung: 4 = hohe Zustimmung) hatten die Schülerinnen und Schüler bei einigen Items auch die Möglichkeit, offen zu antworten. Nach dem ersten Erhebungszeitraum (Schulklasse 1 und 2) wurde das Item SWSF02 (siehe Tabelle 12) verändert und gesplittet, um eine detailliertere Information über die bewusste Schülerwahrnehmung bezüglich Veränderung im Unterrichtsgeschehen zu erhalten.

Tabelle 12: Veränderung des Items V2SWSF02 (alt zu neu)

Item Erhebungszeitraum 1 (Klassen 1 und 2)	Items Erhebungszeitraum 2 (Klassen 3-10)
<ul style="list-style-type: none"> <li>Die Lehrperson hat versucht den Unterricht aufgrund des Feedbacks anzupassen. (vierstufige Likert-Skala)</li> </ul>	<ul style="list-style-type: none"> <li>Sind Dir seit dem Beginn der Feedbacks Veränderungen im Unterricht aufgefallen? (vierstufige Likert-Skala)</li> <li>Falls Du „stimme eher zu“ oder „stimme voll zu“ angekreuzt hast: <ul style="list-style-type: none"> <li>Welche Punkte sind Dir besonders aufgefallen? (offene Antwort)</li> <li>Glaubst Du, dass die Veränderungen durch das Feedback ausgelöst wurden? (vierstufige Likert-Skala)</li> </ul> </li> </ul>

Auswertungsergebnisse, die sich auf diese(s) Item(gruppe) beziehen werden dementsprechend in den Abschnitten 5.2.4.2 und 5.2.4.3 getrennt dargestellt.

#### **5.2.4.1 Items, die in allen 10 Klassen erhoben wurden: Wahrgenommene Einstellung der Lehrperson und Einstellung der Schülerinnen und Schüler gegenüber dem Feedbackverfahren**

Tabelle 13 listet die deskriptive Übersicht der zu Post erhobenen Items mit geschlossenem Antwortformat auf, die in allen 10 Klassen erhoben wurden. Es zeigt sich, dass die Schülerinnen und Schüler den Eindruck hatten, dass die Lehrpersonen das Feedback, welches ihnen gegeben wurde, ernst nahmen und sie empfanden die Möglichkeit, Feedback geben zu können, positiv. Allerdings wird die Häufigkeit der Feedbackrunden (vier Feedbacks in acht bis zehn Wochen) als eher zu hoch eingeschätzt. Gleichzeitig sagte ein beträchtlicher Anteil der Schülerinnen und Schüler aus, dass sie gerne noch häufiger Feedback gegeben hätte. Insgesamt streuten die Einschätzungen relativ stark, was an den Standardabweichungen von ~1 zu erkennen ist.

Tabelle 13: Deskriptive Analyse der Items V2SWSF01, V2SWSF03, V2SWSF04, V2SWSF05, V2SFEE01 und V2SFEE02

	N	AM	SE	SD
Die Lehrperson nahm das Feedback ernst.	210	3,44	0,05	0,73
Ich fand es gut, der Lehrperson Feedback geben zu können.	211	3,58	0,05	0,66
Die Anzahl der Feedbacks in den letzten Wochen war zu hoch.	210	3,22	0,06	0,83
Ich hätte in den letzten Wochen gern noch häufiger Feedback gegeben.	211	2,46	0,08	1,11
Wie oft kam es zur Vorstellung der Ergebnisse durch die Lehrperson?	210	3,17	0,07	1,08
Wie oft wurden die Ergebnisse in der Klasse diskutiert?	207	2,90	0,08	1,14

Die Tabelle 14 zeigt die positiven, Tabelle 15 die negativen Schüleräußerungen zum Feedbackverfahren. Die Aussagen wurden im Auswertungsprozess kategorisiert und die jeweiligen Nennungen aufsummiert. Zunächst fällt auf, dass die Gesamtanzahl der positiven Äußerungen mehr als dreimal so hoch ist wie die der negativen. Dies kann als Hinweis für eine generell positive Einstellung gegenüber dem Verfahren gewertet werden.



Betrachtet man Tabelle 14 genauer, so sehen die Schülerinnen und Schüler die positiven Aspekte vor allem in der Chance, ihre Meinung zum Unterricht äußern zu können und dass ihnen die Möglichkeit gegeben wird, helfen zu können, den Unterricht zu verbessern. In der Quantität nicht so prägnant, jedoch ebenfalls auffällig sind die positiven Äußerungen bezüglich der Anonymität des Verfahrens, der Sichtbarmachung des allgemeinen Stimmungsbilds der Klasse und des Eindrucks, dass die Lehrperson, die von den Schülerinnen und Schüler geäußerte Kritik, umsetzt.

Tabelle 14: Item V2OFRA05 Anzahl der Nennungen

***So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich gut, weil....(max. 3 Stichpunkte)***

Lehrperson helfen den Unterricht zu verbessern	72
Meinung äußern allgemein	50
Allgemeines Stimmungsbild der Klasse sehen	25
Die Lehrperson die Kritik umsetzt	25
Anonymität hilft	21
Negative Kritik äußern können	19
Unterrichtszeit wird weniger	13
Schülerinnen und Schüler werden zur Reflektion des Unterrichts Angeregt	13
Positive Kritik äußern können	12
Kommunikation zwischen Schülerinnen und Schüler und Lehrperson verbessert	7
Gut für das Klassenklima	5
Es hat Spaß gemacht	5
Fand es lustig	3
Abwechslung vom Alltag	2
Sonstiges	17
<b>Gesamt</b>	<b>289</b>

Die Gesamtzahl der negativen Schüleräußerungen zum Feedbackverfahren liegt deutlich tiefer (Tab. 15). Hier geben die Schülerinnen und Schüler vor allem an, dass sie zu wenig, beziehungsweise keine Änderungen im Unterricht wahrgenommen haben, dass zu viel Zeit für das Verfahren aufgewendet werden muss und dass der Feedbackprozess von Mitschülern nicht ernst genommen wird. Zu all diesen Antwortkategorien liegt die Nennungshäufigkeit (meist deutlich) unter 10%.

Tabelle 15: Item V2OFRA06 Anzahl der Nennungen

<b>So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich schlecht, weil....(max. 3 Stichpunkte)</b>	
Zu viel Zeit geht verloren	20
Es hat sich zu wenig geändert	11
Mitschüler nehmen es nicht ernst	10
Es hat sich nichts geändert	5
Zu wenig Zeit zum Ausfüllen	5
Macht keinen Spaß	5
Fragebogen immer gleich	3
Antwortmöglichkeiten zu ungenau	3
Fragen innerhalb des Bogens wiederholten sich	2
Zusammenfassung der Individualurteile zum Klassenurteil	2
Die Ergebnisse fielen immer gleich aus	1
Die Lehrperson nahm es nicht ernst	1
Sonstiges	19
<b>Gesamt</b>	<b>87</b>

Zusammenfassend sehen die Schülerinnen und Schüler das Verfahren durchaus positiv, da ihnen die Möglichkeit gegeben wird, anonym „gehört“ zu werden und den Unterricht positiv mitbeeinflussen zu können. Allerdings sind einzelne Schülerinnen und Schüler eher enttäuscht vom Ausmaß der tatsächlichen Veränderung im Unterrichtsgeschehen oder sie beklagen den hohen Zeitaufwand.

#### **5.2.4.2 Wahrgenommener Einfluss des Feedbacks auf das Lehrerhandeln und den Unterricht Klasse 3-10**

In den Klassen 3-10 wurden die Schülerinnen und Schüler zunächst gefragt, ob sie eine Änderung des Unterrichts während des Erhebungszeitraums festgestellt haben. Dieser Frage stimmten die Schülerinnen und Schüler im Durchschnitt eher nicht zu (Tabelle 16).

Tabelle 16: Deskriptive Analyse Item V2SWSF02 (neu)

	<b>N</b>	<b>Minimum</b>	<b>Maximum</b>	<b>AM</b>	<b>SE</b>	<b>SD</b>
<b>Wahrgenommene Veränderung des Unterrichts</b>	170	1,00	4,00	2,15	0,07	0,91

Die Schülerinnen und Schüler, die eine Änderung wahrgenommen haben (Item-Wert von V2SWSF02 mindestens „3“), wurden gebeten diese stichpunktartig zu beschreiben. Diese Aussagen wurden ebenfalls kategorisiert und die jeweiligen Nennungen gezählt (Tab. 17). Auffällig ist hierbei, dass sich nach Wahrnehmung der Schülerinnen und Schüler vor allem die Unterrichtsatmosphäre verbessert hat, teilweise die Erklärungen der Lehrpersonen verständlicher wurden, die Lehrperson das Unterrichtstempo besser an die Schülerinnen und Schüler angepasst hat und mehr Experimente durchgeführt wurden.

Tabelle 17: Item SWSF21 Anzahl der Nennungen

<b>"Welche Veränderungen im Unterricht sind dir besonders aufgefallen?" (nur berücksichtigt wenn SWSF02 mit mindestens "stimme eher zu" wurde)</b>	
Disziplin/Atmosphäre	18
Besser erklärt	8
Unterrichtstempo/Wiederholungen	6
Mehr Experimente	5
Unterricht ist fokussierter	4
Allgemein bemühter	4
Lehrperson-Schülerinnen und Schüler-Verhältnis verbessert	2
Konkrete Kritikpunkte aus dem Feedback verbessert	2
Schülermitbeteiligung	2
Abwechslungsreicher	1
Struktur	1
Mehr Humor	1
Lehrperson hat sich über Ergebnisse lustig gemacht	1
Sonstiges	16
<b>Gesamt</b>	<b>70</b>

Bei der Folgefrage, ob diese Änderungen auf das Schülerfeedbackverfahren zurückzuführen sind, wurden erneut nur die Antworten der Schülerinnen und Schüler berücksichtigt, die im vorigen Item V2SWSF02 mindestens einen Wert von „3“ angaben. Beachtenswert ist, dass die Schülerinnen und Schüler, die eine Veränderung im Unterricht wahrgenommen haben, dies größtenteils auf das Feedbackverfahren zurückführten (Tab. 18). Der Mittelwert des Items liegt bei 3,16 und keine/r dieser Schülerinnen oder Schüler kreuzte beim Item V2SWSF22 „stimme überhaupt nicht zu“ an.

Tabelle 18: Deskriptive Analyse Item V2SWSF22

	<b>N</b>	<b>Mini- mum</b>	<b>Maxi- mum</b>	<b>AM</b>	<b>SE</b>	<b>SD</b>
<b>Ursache für Veränderung lag im Feedback</b>	51	2	4	3,16	0,09	0,61

Insgesamt sind es etwas weniger als ein Drittel der Schülerinnen und Schüler (vgl. Spalte N der Tabellen 16 und 18), die nach eigener Einschätzung eine Veränderung des Unterrichts wahrgenommen haben. Dies scheint zunächst im Widerspruch zu den verbesserten Skalenwerten der Unterrichtsqualität zu stehen (siehe Abschnitt 5.2.3). Überprüft man jedoch die Korrelation zwischen der bewussten Wahrnehmung („Ich habe Veränderung im Unterricht wahrgenommen“) mit den ermittelten Differenzen der einzelnen Unterrichtsskalen (gemäß Abschnitt 5.2.3), so zeigt sich, dass diese in acht von zwölf Fällen signifikant positiv ausfällt und keine negative Kor-

relation auftritt. Fünf Skalen, die eine signifikante Korrelation ergeben, sind mit den Skalen, die eine signifikant positive Veränderung zu Post erfahren, identisch. Nur eine Skala (*Interessantheit*) zeigt keine signifikante Korrelation, trotz signifikant positiver Veränderung von Prä zu Post. Hieraus lässt sich ableiten, dass ein Zusammenhang besteht, zwischen dem Grad an wahrgenommener Veränderung und der Differenz der beurteilten Unterrichtsqualität zwischen Prä- und Posterhebung.

Tabelle 19: Korrelation zwischen wahrgenommener Veränderung des Unterrichts und Differenz der Unterrichtsqualität (Prä-Post)

Item: Ich habe Veränderungen im Unterricht wahrgenommen.				
		Pearson-Korrelation	Sig. (2-seitig)	N
Differenz der Unterrichtsqualitäts-Skalen (MZ 2 - MZ 1)	Verhältnis zur Lehrperson	,306**	,000	169
	Inhaltliche Strukturiertheit	,177*	,021	169
	Formal kognitive Strukturiertheit	,169*	,028	169
	Angemessenheit Schwierigkeit	,014	,855	169
	Angemessenheit Tempo	,111	,152	169
	Diagnostische Kompetenz	,172*	,026	168
	Klassenmanagement	,208**	,007	170
	Bewertung der Lehrperson	,129	,094	169
	Zeitnutzung	,253**	,001	170
	Interessantheit	,086	,265	169
	Motivierende Unterstützung durch Lehrperson	,266**	,000	170
	Klarheit der Lehrperson	,224**	,004	165

Betrachtet man die bisher vorgestellten Ergebnisse der Schülerfragebogen, so lässt sich zusammenfassend sagen, dass die Beurteilung des Unterrichts zu Post insgesamt positiver ausfällt, dass diese Veränderung von einigen Schülerinnen und Schüler bewusst wahrgenommen wird und dass diese die Gründe hierfür im Feedbackverfahren sehen. Diesem Feedbackverfahren stehen die meisten Schülerinnen und Schüler positiv gegenüber, selbst wenn sie teilweise bewusst keine direkten Auswirkungen auf die Unterrichtsqualität erkennen können.

#### 5.2.4.3 Wahrgenommener Einfluss des Feedbacks auf das Lehrerhandeln in den Klassen 1 und 2

Da, wie oben aufgeführt, das Item V2SWSF02 nach dem ersten Teil der Haupterhebung verändert wurde, wird auf das Ergebnis nur kurz eingegangen. Die Schülerinnen und Schüler der Klassen 1 und 2 nehmen die Bemühungen der Lehrperson, den Unterricht aufgrund des Feedbacks anzupassen als durchschnittlich wahr. Die negativen Beurteilungen in der offenen Frage (Item V2OFRA06) finden sich hier zum Teil wieder. Zu den angepassten Items in den Klassen 3-10 findet sich kein Widerspruch.

### 5.2.5 Verlauf und Tendenz der Schülerkurzfeedbacks über alle Schulklassen

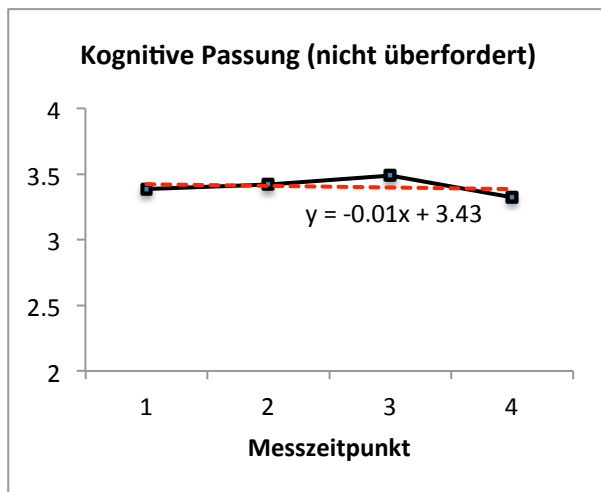


Abbildung 5: Messzeitpunkte 1-4, Verlauf Kognitive Passung (keine Überforderung) über alle Schulklassen mit linearer Regression (rot gestrichelt), Signifikanzniveau  $p$  der Steigung = 0,576

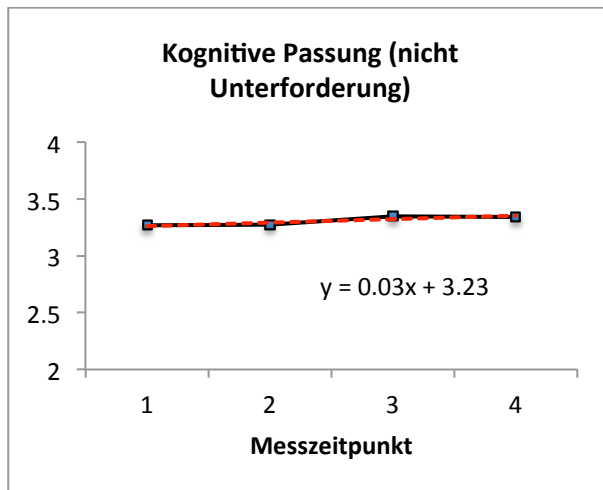


Abbildung 6: Messzeitpunkte 1-4, Verlauf Kognitive Passung (nicht unterfordert) über alle Schulklassen mit linearer Regression (rot gestrichelt), Signifikanzniveau  $p$  der Steigung = 0,118

Neben der Prä-/Posterhebung kann auch der Verlauf der Kurzfeedbacks, also die Daten, die aus der Intervention stammen, interessante Ergebnisse bezüglich der Entwicklung des Unterrichts liefern. Es sei an dieser Stelle nochmals darauf hingewiesen, dass sich die Daten der jeweiligen Messzeitpunkte immer nur auf eine Schulstunde oder Doppelstunde beziehen und dass die Lehrpersonen, abgesehen von der Durchführung der Feedbacks, keine weiteren Vorgaben zur Durchführung der Lektion erhielten. Ziel war es, einen möglichst praxisnahen Datensatz zu erhalten.

Die Abbildungen 5 bis 9 zeigen den durchschnittlichen Verlauf der einzelnen Skalen über die vier Messzeitpunkte und alle Schulklassen. Als Kenngrösse für eine generelle Tendenz im Ergebnisverlauf dient die Steigung der Ausgleichsgerade über alle vier Messzeitpunkte.

Gesamthaft betrachtet fällt auf, dass - mit Ausnahme der Skala *Experimentieren als theoriegeleitetes Handeln* zum Messzeitpunkt 1 - alle Skalen zu allen Messzeitpunkten Werte aufweisen, die über dem mittleren Wert von 2,5 liegen. Besonders hoch sind die Mittelwerte bei den Skalen zur Kognitiven Passung (Abb. 5 und 6), die eine Tendenz zum Deckeneffekt zeigen. Außerdem lassen sich kei-

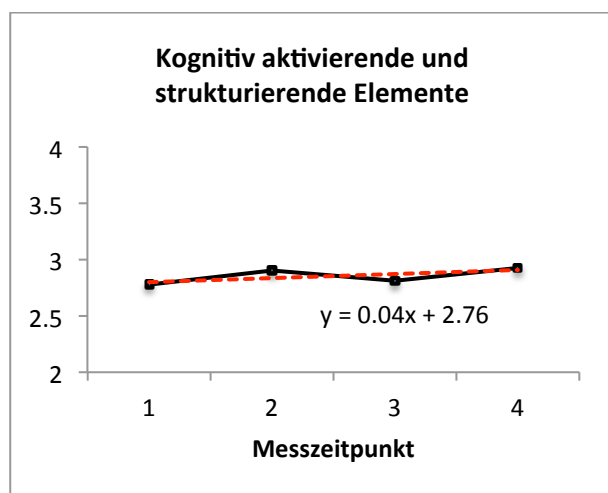


Abbildung 7: Messzeitpunkte 1-4, Verlauf *Kognitiv aktivierende und strukturierende Elemente* über alle Schulklassen mit linearer Regression (rot gestrichelt), Signifikanzniveau  $p$  der Steigung = 0,078

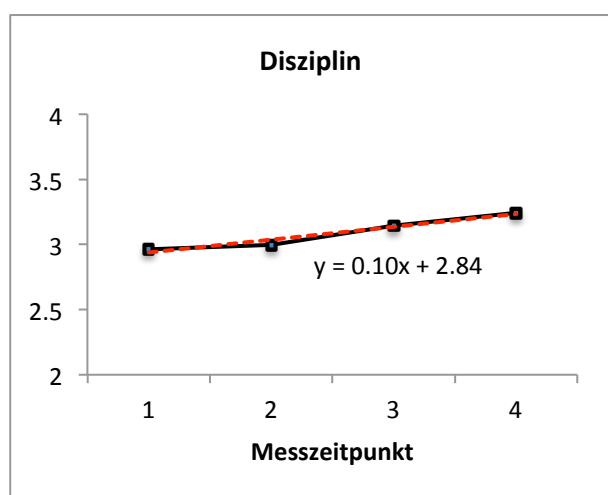


Abbildung 8: Messzeitpunkte 1-4, Verlauf *Disziplin* über alle Schulklassen mit linearer Regression (rot gestrichelt), Signifikanzniveau  $p$  der Steigung = 0,000\*\*\*

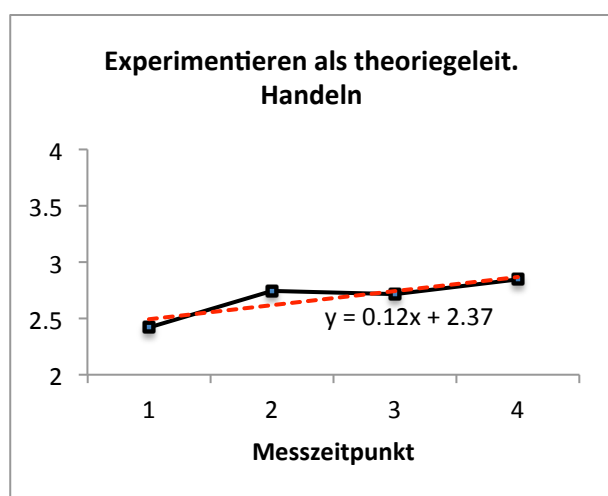


Abbildung 9: Messzeitpunkte 1-4, Verlauf *Experimentieren als theoriegeleitetes Handeln* über alle Schulklassen mit linearer Regression (rot gestrichelt), Signifikanzniveau  $p$  der Steigung = 0,000\*\*\*

ne auffälligen Entwicklungen in beiden Skalen über die vier Messzeitpunkte feststellen (Steigung unterscheidet sich nicht signifikant von 0). Die Skala *Kognitiv aktivierende und strukturierende Elemente* (Abb. 7) weist eine leicht positive Steigungstendenz (marginal signifikant) auf.

Die beiden Skalen *Disziplin* (Abb. 8) und *Experimentieren als theoriegeleitetes Handeln* (Abb. 9) entwickeln sich signifikant positiv.

Aus den Abbildungen wird deutlich, dass die Verläufe über die gesamte Stichprobe durchaus Schwankungen unterliegen. Das hier verwendete Verfahren mit der Steigung der Ausgleichsgeraden als Kenngröße lässt aber den Schluss zu, dass im Zeitraum der Intervention in den beiden Skalen *Disziplin* und *Experimentieren als theoriegeleitetes Handeln* eine positive Entwicklung zu beobachten ist.

Neben den Ankreuzitems konnten die Schülerinnen und Schüler auch in zwei offenen Antwortformaten zur Schulstunde Stellung nehmen (*Folgende Themen des heutigen Unterrichts sollten nochmals angesprochen werden, damit ich sie verstehe und Was ich unbedingt noch zu der heutigen Schulstunde sagen will*). Da diese Aussagen vor allem situations- und klassenspezifisch bedeutsam sind werden sie an dieser Stelle nicht analysiert, jedoch später bei der Charakterisierung einzelner Klassen beziehungsweise von Gruppen im Unterkapitel 5.4 mit einbezogen.

### 5.2.6 Die schriftliche Befragung der Lehrpersonen

Ebenso wie die schriftliche Befragung der Schülerinnen und Schüler war auch diejenige der Lehrpersonen in ein Prä-/Postdesign eingebettet und hatte zum Ziel, Informationen und Entwicklungen zu folgenden drei übergeordneten Fragestellungen zu erhalten:

- Wie schätzen die Lehrpersonen die sozial-kognitiven Voraussetzungen auf Schülerseite, das Schüler/-innen-Lehrpersonen-Verhältnis und die Konsequenzen für ihr eigenes Unterrichtshandeln ein (*Selbsteinschätzung Klassenmanagement, Lehrperson-Schülerinnen und Schüler-Verhältnis, Lernvoraussetzungen der Klasse und leistungsmäßige Homogenität der Klasse*)?
- Welche didaktischen Strategien setzen die Lehrpersonen während des Unterrichts ein?
- Wie schätzen die Lehrpersonen die Schülerinnen und Schüler in ihrer Fähigkeit ein, Unterricht valide bewerten zu können?

Aufgrund der geringen Stichprobenzahl (10) der Lehrpersonen dienen die Daten hauptsächlich zum Gruppenvergleich (siehe Unterkapitel 5.4) und nicht zur quantitativen Betrachtung über die gesamte Stichprobe. Um jedoch einen Überblick zu erhalten, in welchem Bereich sich die jeweiligen Ergebnisse der Befragung bewegen, folgt eine kurze deskriptive Übersicht über beide Erhebungszeitpunkte. Erneut handelte es sich vorwiegend um Items mit geschlossenem Antwortformat, innerhalb dem die Lehrpersonen auf einer vierstufigen Likert-Skala antworten konnten.

#### *Einschätzung der sozial-kognitiven Voraussetzungen*

Betrachtet man die Mittelwerte der vier Skalen, beziehungsweise Items in Tabelle 20, lässt sich generell festhalten, dass die Ergebnisse zu beiden Messzeitpunkten im Durchschnitt sehr ähnlich ausfielen, wobei die Lehrpersonen die Lernvoraussetzungen der Schülerinnen und Schüler und das eigene Klassenmanagement als durchschnittlich, das Schüler-Lehrpersonen-Verhältnis als gut und die Klassen leitungsmäßig eher homogen einschätzten.

Tabelle 20: Einschätzung der Lernvoraussetzungen der Schülerinnen und Schüler, des eigenen Klassenmanagements, der Homogenität und des Schüler-Lehrpersonen-Verhältnisses durch die Lehrpersonen.

	N	Prä		Post	
		AM	SD	AM	SD
<b>Günstige Lernvoraussetzung der Schülerinnen und Schüler</b>	10	2,55	0,42	2,52	0,33
<b>Lehrereinschätzung gutes Klassenmanagement</b>	10	2,80	0,71	2,80	0,72
<b>Lehrereinschätzung gutes Schüler-Lehrpersonen-Verhältnis</b>	10	3,30	0,60	3,20	0,65
<b>Leistungsmäßig Homogen (=1) - Heterogen (=4)</b>	10	1,90	0,74	1,90	0,57

### *Selbstbeschreibung des Unterrichtshandelns*

Tabelle 21 zeigt die Selbsteinschätzung bezüglich der Entwicklung des Unterrichtshandelns zu beiden Messzeitpunkten. Da die Stichprobenzahl von 10 Lehrpersonen gering war und es sich somit als schwierig darstellt, signifikante Unterschiede zwischen der Prä- und Posterhebung festzustellen, wurde versucht die Veränderung des Unterrichtshandelns auf zwei Wegen zu erfassen: So wurden (1.) die Lehrpersonen zur Prä- und Posterhebung gefragt *In welchem Maße sie folgende Strategien im Unterricht einsetzen*, woraufhin die 16 Items der Tabelle 21 folgten, und (2.) wurden die Lehrpersonen bei der Posterhebung zu den gleichen 16 Items gefragt, *wie sich die Häufigkeit der Strategien aufgrund des Schülerfeedbacks verändert hat, im Vergleich zum Unterricht vor der Feedbackphase* (siehe die letzten drei Spalten der Tab. 21 und Anhang 8.1.5, vgl. *konvergenzoriente Daten-Triangulation* Abschnitt 4.5.2).

Betrachtet man die letzten drei Spalten isoliert, gewinnt man den Eindruck, dass die Lehrpersonen an vielen Stellen eine Verhaltensänderung in der Unterrichtsentwicklung gezeigt haben. Dies gilt vor allem für einige Items, die das konstruktivistische Unterrichtshandeln betonen (SKH03, SKH04, SKH06, SKH07, SKH08, SKH09, SKH12, SKH15). Die Beschreibung der Verhaltensänderung wird durch die Mittelwertunterschiede der Verhaltensbeschreibung von Prä- zur Posterhebung allerdings nicht bestätigt: keine Differenz ist signifikant.

Es war zunächst vorgesehen, die hier gewonnenen Daten in die qualitative Analyse der Interviews miteinzuflechten und die dort getätigten Handlungsbeschreibungen zu unterstützen. Allerdings stellte sich heraus, dass die Angaben der Lehrpersonen zum einen oft widersprüchlich sind (bezüglich der Punkte 1 und 2 vorangegangener Absatz) und zum anderen keinen weiteren Erklärungswert besitzen. Die quantitativen Daten werden somit in der Auswertung nicht weiter verwendet.



Tabelle 21: Selbsteinschätzung der eingesetzten didaktischen Unterrichtsstrategien durch Lehrperson, Anzahl der Nennungen über die gesamte Stichprobe

In meinem Unterricht....	Prä			Post			Veränderung zur Zeit vor der Feed-backphase			
	N	AM	SD	AM	SD	Gültig	seltener	gleich	häufiger	
SKH01 lasse ich die Schülerinnen und Schüler selbst Wege zur Lösung von Problemen entdecken	10	2,4	0,52	2,7	0,67	10	0	8	2	
SKH02 lasse ich die Schülerinnen und Schüler ihre eigenen Lösungs-ideen diskutieren.	10	2,6	0,70	3	0,67	10	0	9	1	
SKH03 haben die Schülerinnen und Schüler Gelegenheit in Paaren/Kleingruppen Anwendungsprobleme gemeinsam zu lösen.	10	2,6	0,52	2,8	0,63	10	0	5	5	
SKH04 erfrage ich die Vorstellungen und subjektiven Erklärungen meiner Schülerinnen und Schüler und baue auf diesen auf.	10	3,1	0,32	3,1	0,57	10	0	5	5	
SKH05 erfasse ich den individuellen Fähigkeitsstand und gebe den Schülerinnen gezielt angepasste Aufgaben und Rückmeldung.	10	2,3	0,82	2,2	0,92	10	0	10	0	
SKH06 wähle ich Aufgaben so aus, dass sie den Aufbau und die Erweiterung naturwissenschaftlicher Konzepte ermöglichen.	10	3,3	0,67	3,1	0,74	10	0	6	4	
SKH07 unterstütze ich die Schülerinnen und Schüler dabei, ihr eigenes Lernen wertzuschätzen und Vertrauen in ihr Lernvermögen zu entwickeln.	10	3,5	0,53	3,4	0,52	10	0	6	4	
SKH08 unterstütze ich das forschend-entdeckende Lernen der Schülerinnen und Schüler durch unterschiedliche Methoden und Lernarrangements	10	2,4	0,52	2,6	0,52	10	1	3	6	
SKH09 lasse ich die Schülerinnen und Schüler neues Wissen durch Experimentieren entwickeln.	10	2,4	0,52	2,4	0,52	10	0	6	4	
SKH10 gebe ich den Schülerinnen und Schüler die Möglichkeit, aus Fehlern oder falschen Vorstellungen zu lernen.	10	3,2	0,63	3,2	0,63	9	0	7	2	
SKH11 erkläre ich neue Inhalte und Zusammenhänge ausführlich.	10	3,6	0,52	3,6	0,52	10	0	7	3	
SKH12 gebe ich genaue Anweisungen, wie die Aufgaben zu erledigen sind.	10	3,4	0,70	3,2	0,79	10	0	6	4	
SKH13 achte ich besonders darauf, dass das richtige Fachvokabular benutzt wird.	10	3,2	0,63	3,1	0,74	10	0	8	2	
SKH14 ist es mir wichtig, dass die Schülerinnen und Schüler den durchgenommenen Stoff exakt wiedergeben.	10	2,6	0,70	2,4	0,97	10	0	9	1	
SKH15 finden ausführliche Klassengespräche statt.	10	2,9	0,57	2,6	0,84	10	1	5	4	
SKH16 ist es mir sehr wichtig mit dem Stoff durchzukommen.	9	2,1	0,93	2,4	0,84	10	1	9	0	

### *Einschätzung des Unterricht-Beurteilungsvermögens der Schülerinnen und Schüler*

Die Lehrpersonen schätzten das Beurteilungsvermögen der Schülerinnen und Schüler überdurchschnittlich ein. Dies ändert sich auch zur Posterhebung kaum, wenngleich eine leicht negative Tendenz festzustellen ist, die jedoch aufgrund der geringen Stichprobengröße keine statistische Aussagekraft besitzt.

Tabelle 22: Lehrereinschätzung: Beurteilungsvermögens von Schülerinnen und Schüler bezüglich Unterricht

	Prä			Post	
	N	MW	SD.	MW	SD
<b>Lehrereinschätzung Beurteilungsvermögen Schülerfeedback</b>	10	3,08	,45	2,92	,52

Die Lehrpersonen scheinen demnach in der Gesamtbetrachtung nicht in ihrer Einschätzung durch das Feedbackverfahren beeinflusst worden zu sein und schätzen das Beurteilungsvermögen der Schülerinnen und Schüler, sowohl vor, als auch nach der Intervention generell positiv ein.

### **5.2.7 Die Interviews mit den Lehrpersonen**

Im Rahmen der leitfragengestützten Interviews wurde im Wesentlichen auf die drei übergeordneten Aspekte der Forschungsfragen eingegangen:

1. Auf die Mechanismen des Schülerfeedbacks während der Intervention, also die Reaktionen und Verhaltensweisen der Lehrpersonen auf Grundlage des Rahmenmodells der Unterrichtsentwicklung nach Helmke und Hosenfeld (2005), welche sich sowohl auf das Schüler-Kurzfeedback im Allgemeinen als auch auf explizite Ergebnisse des selbigen bezogen.
2. Auf die Wirkung des Schülerfeedbackverfahrens auf die nicht-leistungsbezogene Einschätzung der Lehrpersonen bezüglich der Schülerinnen und Schüler (Reife, Beurteilungsvermögen von Unterricht, wahrgenommene Wertschätzung der eigenen Arbeit, Disziplin und Schüler-Lehrpersonen-Verhältnis).
3. Auf die eingeschätzte Praktikabilität des Feedbackverfahrens.

Der erstgenannte Punkt unterscheidet sich vom zweiten und dritten dahingehend, dass dieser keine abschließende Bewertung vornimmt, sondern die Entwicklung während des Verfahrens abfragt. Somit können die hier genannten Punkte wichtige Indizien für die Wirkung von Schülerfeedback auf die Unterrichtsqualität in den einzelnen Schulklassen liefern. Deshalb erfolgt eine detaillierte Vorstellung dieser Interviewergebnisse in der gruppenweisen Auswertung in Unterkapitel 5.5. An dieser Stelle werden im Sinne einer deskriptiven Übersicht lediglich die Häufigkeiten der genannten Aussagen gezählt und in Tabelle 23 aufgeführt.

Die beiden letztgenannten Aspekte bilden eine abschließende Bewertung des Verfahrens und dessen Wirkung und werden somit in den folgenden Abschnitten im Detail vorgestellt.

Da Mehrfachnennungen in den Interviews erlaubt waren, beziehungsweise einige Aspekte multikausal erklärt wurden, ist in den dazugehörenden Tabellen die Gesamtzahl der Nennungen der einzelnen Aspekte und deren Ausprägungen meist höher als der Umfang der Stichprobe ( $n = 10$ ).

#### **5.2.7.1 Unterrichtsentwicklungsbezogene Reaktionen und Verhaltensweisen der Lehrpersonen auf das Schülerfeedback**

Über die gesamte Stichprobe gesehen, nahmen die Lehrpersonen die Ergebnisse der Schülerfeedbacks überwiegend nüchtern auf (siehe Tab. 23) und stimmten den Aussagen der Schülerinnen und Schüler zu. Allerdings sahen die Lehrpersonen in über der Hälfte der Fälle keine Änderungsmöglichkeit der geäußerten Kritikpunkte. Sie attribuierten also häufig auf unkontrollierbare Faktoren. Sahen sie jedoch eine Möglichkeit, die kritischen Punkte zu verbessern, bildete sich daraus nach eigener Einschätzung vielfach eine Handlungsabsicht und eine tatsächliche Veränderung der Handlung. Außerdem wurde mehrfach geäußert, dass die Lehrpersonen durch die Feedbacks motiviert wurden, zukünftig den Unterricht aufgrund der Ergebnisse, zu verändern.

Tabelle 23: Interview; Verhalten im Rahmenmodell der Unterrichtsentwicklung (allgemein und speziell); Anzahl der Nennungen über die gesamte Stichprobe

<b>Reaktionen auf das Schülerfeedback im Rahmen der Unterrichtsentwicklung; <i>allgemein</i></b>		<b>Anzahl der Nennungen</b>
<b>Rezeption</b>	emotional positiv	16
	emotional negativ	11
	nüchtern	49
<b>Interpretation</b>	Schüleraussage richtig aber nicht veränderbar	32
	Schüleraussage richtig und veränderbar	24
	Schüleraussage falsch	6
<b>Motivation</b>	Änderungsabsicht vorhanden	39
	keine Änderungsabsicht	13
<b>Handlung</b>	Handlung verändert	32
	Handlung unverändert	12
<b>Motivation für zukünftiges Unterrichtshandeln</b>	Beschreibung der Handlungsabsicht	13

Insgesamt betrachtet lässt sich die Tendenz ablesen, dass Lehrpersonen aufgrund des Schülerfeedbacks angeregt werden, kritische Punkte in ihrem Unterricht zu verändern, sofern sie diese für kontrollierbar halten.

### 5.2.7.2 Lehrer-Einschätzungen zu sozial-kognitiven Voraussetzungen und ihre Entwicklungen in der Feedbackphase

Wie bereits oben erwähnt, bezieht sich die Art und Weise, wie Lehrpersonen die Schülerinnen und Schüler einschätzen auf fünf Fragestellungen:

- Schätzen Lehrpersonen das Beurteilungsvermögen der Schülerinnen und Schüler nach der Intervention höher ein?
- Schätzen die Lehrpersonen die Schülerinnen und Schüler nach der Intervention als intellektuell reifer ein?
- Nehmen die Lehrpersonen eine höhere Wertschätzung der eigenen Arbeit durch die Schülerinnen und Schüler wahr?
- Nehmen die Lehrpersonen die Schülerinnen und Schüler als disziplinierter wahr?
- Hat sich das Schüler-Lehrpersonen-Verhältnis insgesamt verbessert?

Tabelle 24 zeigt eine Übersicht der Kategorien, gruppiert nach den Fragestellungen, mit der dazugehörigen Anzahl der Nennungen.

Tabelle 24: Interview; Entwicklung, wie Lehrpersonen die Schülerinnen und Schüler wahrnehmen

		Anzahl der Nennungen
<b>Einschätzung des Beurteilungsvermögens der Schülerinnen und Schüler bzgl. Unterricht</b>	gestiegen	2
	gleich geblieben	8
	gesunken	1
<b>Allgemeine Einschätzung der Schülerinnen und Schüler bzgl. der Reife</b>	gestiegen	5
	gleich geblieben	7
	gesunken	0
<b>Wahrgenommene Wertschätzung der eigenen Arbeit durch die Schülerinnen und Schüler</b>	gestiegen	4
	gleich geblieben	8
	gesunken	0
<b>Allgemeine Einschätzung des Schüler-Lehrpersonen-Verhältnisses</b>	verbessert	10
	gleich geblieben	5
	verschlechtert	1
<b>Disziplin der Schülerinnen und Schüler</b>	verbessert	5
	gleich geblieben	8
	verschlechtert	0

Im Folgenden werden die teilweise wahrgenommenen Entwicklungen eingehender charakterisiert und mit ausgewählten Interviewauszügen belegt.

#### *Einschätzung des Beurteilungsvermögens der Schülerinnen und Schüler bezüglich des Unterrichts*

Hier zeigte sich wie schon bei der schriftlichen Befragung kaum eine Entwicklung. Nur drei Nennungen beschreiben eine Veränderung der Einschätzung (zwei davon positiv), die jedoch während des Interviews unter anderem aufgrund der eingeschätzten Differenzierungsfähigkeit der Schülerinnen und Schüler relativiert werden (I = Interviewer, B = Befragte/r):

Lehrperson Klasse 7:

*I Nehmen Sie denn die Einschätzung der Schülerinnen und Schüler bezüglich der Unterrichtsbewertung ernster?*

*B Prinzipiell ja, ich eh weiß aber nicht, ob sie sich wirklich von Stunde zu Stunde neu Gedanken gemacht haben, das erschließt sich mir jetzt einfach aus den Ergebnissen nicht. Ich habe da finde ich, schon Variationen drin gehabt, das ist aber immer sehr ähnlich gewesen. Das ist wo ich ein bisschen ein Problem sehen würde. Aber ich kann mich auch täuschen und die Schüler machen das eh, absolut präzise.“*

Die zum Teil sachfernen Kommentare der Schülerinnen und Schüler im offenen Kommentarbereich des Schüler-Kurzfeedbacks (siehe Abschnitt 5.4.1.7) scheinen auf die Bewertung des Beurteilungsvermögens kaum Einfluss zu haben. Sie wurden in den Interviews nie hiermit in Zusammenhang gebracht.

#### *Allgemeine Einschätzung der Schülerinnen und Schüler bezüglich der intellektuellen Reife*

Insgesamt vier Lehrpersonen (fünf Nennungen) schätzten die Schülerinnen und Schüler nach dem Feedbackverfahren als reifer ein. Zum einen wird dies u. a. damit begründet, dass die Lehrpersonen positiv wahrnahmen, mit welcher Ernsthaftigkeit die Schülerinnen und Schüler sich am Feedbackverfahren beteiligten und zum anderen, dass sich die Schülerinnen und Schüler bei der Verbesserung der, durch das Feedbackverfahren aufgezeigten Mängel, aktiv beteiligten:

Lehrperson Klasse 5:

*B ...also ich muss sagen, während dem Feedback, ja war dann spätestens mal klar, wie die Sache abläuft. Ich hab tatsächlich auch den Raum verlassen, ich habe, die waren ruhig, die haben gemerkt okay da geht es jetzt um was, wir nehmen das ernst, die haben die selbstständig dann eingesammelt und auch nach vorne gelegt und so weiter, ehm von dem her denke ich hat es, da schon einen positiven Effekt gehabt, ja in dem Sinne, während des Feedbacks...”*

Lehrperson Klasse 6:

*I mhm (<bejahend>), hat sich den das Bild, dass Sie von ihren Schülerinnen und Schülern haben, verändert, also nehmen Sie sie eventuell reifer, erwachsener, oder reflektierter war als vorher?*

*B hmm, doch ein Stückweit schon, doch, muss ich sagen,*

[...]

**I** Okay, eben das, dies Reflektiertheit oder die gewachsene Reflektiertheit, haben Sie dafür jetzt ein konkretes Beispiel? An was Sie das jetzt festmachen? Dass Sie das jetzt, ja an einem Schüler wahrgenommen haben, dass Sie das jetzt vorher noch nicht wahrgenommen haben oder so?

**B** [...] Also mir fällt grad kein konkretes Beispiel ein. Aber was mir generell auffällt, auch in der Phase, dass Schüler doch immer mehr, ehm aufpassen und einen auch festnageln, wenn sie eben merken; doch ich kann ein Beispiel machen: Als es im (<unverständlich>) ums Experimentieren geht und ich an einem Experiment gezeigt habe, was es heißt, wissenschaftlich korrekt zu arbeiten, dann haben die Schüler doch schon genau gesagt ja, aber jetzt ehm, müsste es doch des und des oder ja jetzt zumindest aufgefallen was sie meinen, aber die passen schon mehr auf und versuchen das auch aufzunehmen, was ich denen dann zeige [...]

#### *Wahrgenommene Wertschätzung der eigenen Arbeit*

Zwar sehen vier Lehrpersonen eine gesteigerte Wertschätzung der eigenen Arbeit durch die Schülerinnen und Schüler, allerdings handelt es sich hierbei tendenziell um subjektive Einschätzungen und es können kaum Belege aufgeführt werden. Außerdem wird diese Entwicklung nur einer beschränkten Anzahl von Schülerinnen und Schüler zugeschrieben:

Lehrperson Klasse 3:

**I** Haben Sie denn das Gefühl, dass die Schülerinnen und Schüler die Arbeit der Lehrperson mehr schätzen als vorher?

**B** Diejenigen, die es interessiert schon, ja [...] Bei vielen ist es auch, vielen ist es auch egal.

**I** Mhm (<bejahend>) woran machen Sie das fest?

**B** Na, an Fragen, die die stellen, also es werden manchmal so Fragen gestellt, könnte man das jetzt auch anders machen, eh, kann man das nicht auch ausprobieren. Und ja da muss ich manchmal sagen, ja das geht jetzt grad nicht oder manchmal machen wir das dann auch und ehm das, das gibt es schon, auch jetzt die Wertschätzung, meine Arbeit da sich geändert hat, das weiß ich eigentlich nicht so richtig. Es, ich habe das Gefühl, dass einige aufmerksamer sind.

Eine Veränderung der wahrgenommenen Wertschätzung der eigenen Arbeit ist also nur in geringem Maße festzustellen und scheint vernachlässigbar zu sein.

### *Entwicklung der Disziplin der Schülerinnen und Schüler*

Vier Lehrpersonen (fünf Nennungen) sehen eine positive Entwicklung in der Disziplin der Schülerinnen und Schüler, wobei sich die Begründungen unterscheiden. Die Lehrperson der Klasse 3 führt dies auf eine intensive pädagogische Arbeit mit der Schulklasse zurück, die nicht aufgrund der Schüler-Feedbacks sondern aufgrund allgemeiner sozialer Probleme stattfand. Die Lehrperson der Klasse 1 sieht die Ursache in einer verbesserten sozialen Kontrolle innerhalb der Schulklasse (die Schülerinnen und Schüler sorgten gegenseitig für Disziplin) und lediglich die Lehrpersonen der Klassen 5 und 10 identifizierten ihr eigenes Handeln aufgrund der Rückmeldungen als Ursache für die Verbesserung. Gemeinsam war allen drei Klassen (1, 5 und 10), dass sich einige Schülerinnen und Schüler über einen zu hohen Lautstärkepegel beklagt haben. Die Aussage der Lehrperson aus Klasse 5 soll hier als Beispiel dienen:

Lehrperson Klasse 5:

**B** Also das, ehm nach der ersten Erhebung kam eben eine Beschwerde, praktisch, dass oder, dass es zu laut sei im Unterricht und daraufhin habe ich die mehr Ruhe und Konzentration wieder ehm verlangt einfach, und ich denke, ja das kann man dann doch, so eine gewisse Tendenz dann erklären ja.

**I** Okay, damit haben Sie im Prinzip schon die nächste, meine nächste Frage beantwortet. Und ehm bei Feedback 2 wurde dann auch gleich sofort gesagt, ehm ja, das ist, war jetzt schon viel besser, ehm wie haben Sie darauf reagiert, dass die jetzt, eh, dass das verbessert wurde? Die Geschichte mit der Atmosphäre, dass es zu laut sei?

**B** Ja weil ich da direkt im Prinzip, ehm darauf reagiert habe und das direkt auch angesprochen habe und gesagt habe, eh wir müssen wieder eh ne gewisse Arbeitsatmosphäre reinbringen, dass sozusagen diese legere Art, was ich als Ziel hatte, dass man da vielleicht über so eine gewisse Lockerheit reinkommt, nicht passt ja.

Die Information, die das Schülerfeedback über die Arbeitsatmosphäre liefert, scheint für einige Lehrpersonen hilfreich zu sein, da es ihnen vermutlich leichter fällt, für eine angemessene Ruhe in der Klasse zu sorgen, wenn sie wissen, dass einige Schülerinnen und Schüler die Lautstärke ebenfalls als störend empfinden.

### *Schüler-Lehrpersonen-Verhältnis*

Während die schriftliche Befragung kaum eine Entwicklung des Schüler-Lehrpersonen-Verhältnisses erkennen ließ, äußerten sich die Lehrpersonen im Laufe des Interviews häufig positiv zum Einfluss des Schüler-Kurzfeedback-Verfahrens auf das Schüler-Lehrpersonen-Verhältnis.

Nur in einem Falle (Lehrperson 10) wird eine (eingeschränkte) Verschlechterung durch das Schülerfeedback beschrieben (siehe unten). Die meisten Lehrpersonen sahen die Ursachen im Feedbackverfahren; die Interpretation der Mechanismen unterschied sich jedoch von Lehrperson zu Lehrperson. Während beispielsweise die Lehrperson aus Klasse 10 die Selbstreflexion des eigenen Verhaltens der Schülerinnen und Schüler heraushob, sah die Lehrperson aus Klasse 5 die Ursache in der wahrgenommen Veränderung des Unterrichts durch die Schülerinnen und Schüler. Die Lehrperson aus Klasse 1 beschrieb die gemeinsame Reflexion über den Unterricht und die Wertschätzung der Schüleraussagen als Grund für die Verbesserung des Verhältnisses:

Lehrperson Klasse 10:

**B** *Ehm, manche Schüler waren sich dann doch eher bewusst, oder ist es auch selbst zu Bewusstsein gekommen, dass sie da manchmal ein bisschen zu lebhaft sind. Und, da sie vielleicht dem Lehrer auch dadurch das Leben schwerer machen, als es sein müsste, ja, das hat sich dann schon bei manchen wirklich auch so ein bisschen gebessert, weil die haben dann die Frage gelesen und so weiter und so und da hat man schon gemerkt, dass sich da was tut bei anderen, ja...*

Lehrperson Klasse 5:

**B** *Ich denke, dadurch dass ich eine Rückmeldung bekommen hab und einiges für mich auch klarer wurde, was ich jetzt auch gar nicht damit gerechnet habe, ja dass ich auf gewisse Wünsche auch eingehe und gewisse Dinge dann auch beachte, denke ich hat sich das auch verbessert, ja. Dass, dass sie sehen, okay, wir haben das jetzt gesagt, das ist zu laut und die nächste Stunde wird das eingefordert. Das ist dann doch, sagen wir mal, ein bisschen professioneller wurde, ja, also dass die Schüler dann doch das eine oder andere mal, habe ich den Eindruck, ja...*

Lehrperson Klasse 1:

**B** *...Es hat, insofern etwas intensiviert, also, also gemeinsam über Unterricht nachgedacht haben [...] aber ich glaub, und ich glaub das ist das entscheidende, also ich glaub sie fühlen sich schon mehr ernst genommen...*



Lehrperson 10 argumentierte aber auch gleichzeitig für eine Verschlechterung des Schüler-Lehrpersonen-Verhältnisses, was sie vor allem an der strengeren Durchsetzung von Lerndisziplin festmachte:

Lehrperson Klasse 10 :

*I Weil Sie haben vorher mal gesagt, Sie greifen früher und konsequenter ein oder durch. [...]*

*B Vielleicht ist es das, vielleicht ist es dann wirklich für die eh, für die anderen Schüler dann angenehmer, wenn dann am Anfang, eh gleich, bei den entsprechenden Schülern gehandelt wird. [...] Aber ich fand die die Atmosphäre eher nicht mehr so, eh, persönlich oder so eine Beziehung sondern es war doch eher eh klassischer Unterricht mit strenger Disziplin und so...*

Abweichend zur schriftlichen Befragung der Lehrpersonen, die keine Entwicklung des Schüler-Lehrpersonen-Verhältnisses aufzeigte, wird diese in den Interviews als eher positiv dargestellt. Dies könnte darin begründet sein, dass das Verhältnis bereits zu Prä als positiv bewertet wurde und die positiven Veränderungen nicht groß genug ausfielen, um sich in einer vierstufigen Likert-Skala abzubilden. Ein weiterer Grund könnte in der sozialen Erwünschtheit gegenüber dem Interviewer liegen. Da der Interviewer davon ausgeht, dass die Forscher eine möglichst positive Entwicklung erwarten, könnte der/die Befragte tendenziell in diese Richtung antworten. Ähnlichkeit besteht allerdings mit der schriftlichen Befragung der Schülerinnen und Schüler. Diese bewerten das Schüler-Lehrpersonen-Verhältnis nach der Feedbackphase ebenfalls signifikant besser als davor.

Insgesamt bleibt festzuhalten, dass aus Sicht der Lehrpersonen ein Schülerfeedback unterschiedliche Mechanismen in Gang setzen kann, die zu einem verbesserten Schüler-Lehrpersonen-Verhältnis und zu verbesserten sozial-kognitiven Voraussetzungen beitragen. Diese Mechanismen können unter anderem in einer verbesserten Disziplin, einer größeren Wertschätzung der eigenen Arbeit und in einer wahrgenommenen fortgeschrittenen intellektuellen Reife der Schülerinnen und Schüler begründet sein.

### **5.2.7.3 Einschätzungen der Lehrpersonen bezüglich des Feedbackverfahrens**

Ähnlich wie bei den Schülerinnen und Schüler fiel die Beurteilung des Verfahrens insgesamt positiv aus. Wie Tabelle 25 zeigt, überwiegt die Anzahl der positiven Nennungen in allen Kategorien. Auch hier waren Mehrfachnennungen und multikausale Ursachenzuschreibungen erlaubt.

Tabelle 25: Einschätzung des Schüler-Feedbackverfahrens

Kategorien		Anzahl der Nennungen
Das Verfahren insgesamt	gut	29
	problematisch	17
	schlecht	3
Einschätzung der Fragebögen	gut	34
	schlecht	16
Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	21
	schlecht	9
In elektronischer Form sinnvoll	ja	11
	nein	4
Würde es selbst in elektronischer Form einsetzen	ja	9
	nein	1

### *Einschätzung des Verfahrens insgesamt*

Die Lehrpersonen schätzen an dem Verfahren vor allem, dass es zeitökonomisch und problemlos durchführbar ist, dass es den Schülerinnen und Schülern das Gefühl gibt, gehört und ernst genommen zu werden und es zur Kommunikation anregt. Außerdem sehen sie sowohl auf Schüler- und Lehrpersonenseite eine Anregung zum Nachdenken über Unterricht.

Problematisch sehen wenige Lehrpersonen den hohen Zeitaufwand (die Wahrnehmungen diesbezüglich variieren von Lehrperson zu Lehrperson) und dass bei einigen Schülerinnen und Schülern zum Ende der Intervention die Motivation nachließ, die Fragebogen auszufüllen (auch dies wurde von anderen Lehrpersonen gegenteilig wahrgenommen). Einige Lehrpersonen vermissten Vergleichswerte und/oder didaktische beziehungsweise pädagogische Handlungsanweisungen für die entsprechenden Ergebnisse und taten sich deshalb schwer, diese zu interpretieren. Dies wird auch damit begründet, dass die Variationsbreite der Ergebnisse zum Teil relativ klein war. Eine Lehrperson forderte mehr offene Antwortformate, andere wiederum sehen diese als problematisch an, da sich mit der Zeit die sachfernen Kommentare innerhalb dieses Formats häuften.

Folgende Punkte wurden als schlecht kategorisiert, weil sie zentrale Inhalte des Schüler-Kurzfeedbackverfahrens kritisieren: Eine Lehrperson empfand die Feedbackbögen als zu eindimensional und eine andere die Darstellung der Ergebnisse in Diagrammen als zu anspruchsvoll für die Schülerinnen und Schüler. Eine weitere Lehrperson sieht einen zu geringen Fokus in der Anleitung, wie man die Schülerinnen und Schüler in der Diskussion zu Beteiligten macht.

### *Einschätzung der Schüler-Kurzfeedback-Bögen*

Eine häufig getätigte positive Aussage bezüglich der Fragebögen war, ähnlich wie bei den Schülerinnen und Schülern, die Tatsache, dass dieser anonym gehalten ist und dass dieser sowohl die Lehrperson als auch jede/n Schülerinnen und Schüler dazu anregte, über Unterricht, auch aus einer neuen Perspektive, nachzudenken. Des Weiteren äußerten einige Lehrperso-

nen, dass die Skalen hilfreich und sinnvoll waren, sowohl als Anhaltspunkt für die Diskussion, als auch für das eigene Bestreben, den Unterricht zu verbessern. Für drei Lehrpersonen sind die Fragebogenergebnisse für ihre Unterrichtsentwicklung entscheidender als die anschließende Ergebnisvorstellung, beziehungsweise die Diskussion darüber. Ebenfalls dreimal wurde das offene Antwortformat positiv herausgehoben. Eine Lehrperson gab an, dass die Ergebnisse der Fragebögen ihr Sicherheit bei der weiteren Unterrichtsentwicklung gab. Eine andere sah für die Schülerinnen und Schüler einen Vorteil darin, dass diese Feedback geben können, ohne direkte verbale oder nonverbale Rückmeldung durch die Lehrperson zu erhalten (Mimik, Gestik, Kommentar).

Negativ wurde von einigen Lehrpersonen vor allem der benötigte Zeitaufwand beurteilt, der höher war als zunächst vermutet. Vor allem im Hinblick darauf, dass sich das Feedback auf die jeweilige – nun noch kürzere – Schulstunde beziehungsweise Doppelstunde bezog. Eine Lehrperson schlug deshalb vor, nicht alle Skalen auf einmal abzufragen, sondern in einem alternierenden Rhythmus. Außerdem bemängelten zwei Lehrpersonen, dass die gebildeten Skalen für die Diskussion nicht besonders geeignet waren, weil sie entweder zu ungenau beschrieben, welche konkreten Defizite auftraten oder weil die Unterschiede zwischen den jeweils beurteilten Stunden zu gering waren und es somit schwierig war, Auffälligkeiten zu identifizieren. Eine Lehrperson forderte deshalb, dass neben den Skalen auch die Einzelitems in der bereitgestellten Auswertung sichtbar sein sollten. Außerdem wurde unter anderem in Frage gestellt, ob Schülerinnen und Schüler die eigene Denkaktivität (kognitive Aktivierung) adäquat beurteilen können.

### *Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse*

Die Lehrpersonen erachteten die Vorstellung der Ergebnisse als gut und wichtig, vor allem weil einerseits die Schülerinnen und Schüler das Bedürfnis haben, die Ergebnisse zu erfahren und andererseits die Wertschätzung der Schülermeinung verdeutlicht wird. Außerdem wurde von mehreren Lehrpersonen eine Erweiterung der Perspektive auf vielfache Weise beschrieben: Die Schülerinnen und Schüler sehen nicht nur, wie ihre eigene Beurteilung im Vergleich zur Klassenmeinung ausfällt, sondern eventuell auch, wie ihr eigenes Verhalten bei den Mitschülern ankommt. Die Lehrpersonen bekommen, zu den in abstrahierter Form (Mittelwert, Standardabweichung) gehaltenen Fragebogenergebnissen, konkretere Anhaltspunkte für eine Verbesserung des Unterrichts.

Zwei Lehrpersonen beklagten, dass das Interesse der Schülerinnen und Schüler an den Fragebogenergebnissen eher niedrig ausfiel und daher die Diskussionen träge verliefen. Eine weitere stellte fest, dass sich nur ein Teil der Schülerinnen und Schüler an den Diskussionen beteiligte. Außerdem wurde von zwei Lehrpersonen darauf hingewiesen, dass während der Diskussion die Anonymität der Fragebögen, zumindest teilweise, verloren geht. Bezüglich des Faktors Zeit

äußerten sich ebenfalls drei Lehrpersonen negativ: Eine empfand den Zeitaufwand der Diskussion zu hoch, eine andere kritisierte, dass die Abstände zwischen Feedback und Besprechung teilweise zu lang ausfielen – was eine Diskussion über die Ergebnisse erschwerte – und eine dritte bezweifelte die Sinnhaftigkeit eines Zwei-Wochen-Rhythmus im normalen Schulalltag. Ansonsten merkte eine Lehrperson an, dass die Schülerinnen und Schüler sich schwer taten, die Standardabweichung bei der Ergebnisdarstellung zu verstehen.

### *Einsatz des Verfahrens in elektronischer Form*

Zum Schluss des Interviews wurden die Lehrpersonen gefragt, ob sie sich vorstellen könnten, ein elektronisches Programm einzusetzen, dass im Kern dem durchgeführten Verfahren gleicht, die Daten allerdings automatisiert auswertet: Die Schülerinnen und Schüler könnten die Items am Ende der Stunde oder als Hausaufgabe per Smartphone/PC/Tablet beantworten und die Lehrperson bekommt die Daten anonymisiert auf sein EDV-Gerät geliefert.

Alle Lehrpersonen empfanden diese Form des Schülerfeedbacks als sinnvoll und nahezu alle würden es auch selbst in dieser Form einsetzen, gaben aber teilweise zu bedenken, dass dies nicht unter Zwang geschehen dürfe, die Gefahr des Datenmissbrauchs beachtet werden muss und dass eventuell der Rücklauf der Daten gering ausfallen könnte. Eine Lehrperson lehnt den Einsatz des Verfahrens in elektronischer Form eher ab, weil sie aufgrund ihrer bisherigen Erfahrungen die Gefahr sieht, dass auch gute Instrumente oft sinnentleert eingesetzt werden. Soll heißen, man erhebt Daten und gibt sich damit zufrieden, ohne auf die Ursachen der Daten genauer einzugehen.

Auf die Frage in welchem Intervall die Lehrpersonen ein elektronisches Schülerfeedbackverfahren einsetzen würden, reichten die Angaben von spontan über wöchentlich bis ein/zwei mal pro Jahr.

## **5.2.8 Zusammenfassung der Ergebnisse über die gesamte Stichprobe**

Die bisher vorgestellten Ergebnisse über die gesamte Stichprobe geben bereits Anlass für eine vorläufige Beantwortung der Forschungsfragen 1, 2 und 4.

*Lassen sich Hinweise identifizieren bezüglich des Einflusses von Schüler-Kurzfeedback auf a) ausgewählte Skalen der Unterrichtsqualität unter besonderer Berücksichtigung des Schüler-Lehrpersonen-Verhältnisses und b) die Unterrichtsentwicklung und das Unterrichtshandeln der Lehrperson? (FF 1)*

Aus Schülerinnen und Schüler-Sicht zeigen sieben von zwölf Skalen der Unterrichtsqualität eine signifikante Verbesserung von Prä zu Post. Allerdings zeigen nur vier Skalen einen Effekt von  $d > 0.2$ .

Die Lehrpersonen nahmen die Aussagen der Schülerinnen und Schüler ernst und versuchten, sofern sie eine Veränderungsmöglichkeit sahen (also internal kontrollierbar attribuieren), das Unterrichtshandeln entsprechend anzupassen.

Die Daten bezüglich der Selbstbeschreibung des Unterrichtshandelns durch die Lehrperson mit Hilfe geschlossener Items sind relativ diffus und lassen sich nur schwer interpretieren. Eine Aussage über deren Entwicklung ist somit kaum möglich. Eine eingehende Analyse der Interviews wird diesbezüglich weitere wertvolle Daten liefern (siehe Unterkapitel 5.5).

Sowohl auf Schüler- als auch Lehrpersonenseite wurde eine Verbesserung des Schüler-Lehrpersonen-Verhältnisses festgestellt (Lehrpersonen: Bezogen auf die Daten aus den Interviews). Der Effekt über die gesamte Stichprobe auf Schülerseite ist allerdings mit  $d = 0.14$  sehr klein.

*Welche Entwicklungen nehmen Schülerinnen und Schüler im Unterricht innerhalb des Schüler-Kurzfeedback-Verfahrens wahr? (FF 2)*

Die Schülerinnen und Schüler nahmen während der Feedbackphase Entwicklungen im Unterricht eher selten bewusst wahr. Falls doch, führten sie diese meist auf das Feedbackverfahren zurück. Besonders häufig stellten sie dabei eine Verbesserung der Unterrichts Atmosphäre und eine gesteigerte Disziplin fest. Außerdem nahmen sie eine erhöhte Anzahl durchgeführter Experimente und eine größere Bemühung auf Seiten der Lehrperson wahr, den Unterrichtsstoff besser zu erklären und das Unterrichtstempo anzupassen.

*Wie schätzen Lehrpersonen und Schülerinnen und Schüler die Praktikabilität des Schüler-Kurzfeedback-Verfahrens ein? (FF4)*

Das Schülerfeedbackverfahren wurde sowohl auf Schülerinnen und Schüler- als auch Lehrpersonenseite überwiegend positiv beurteilt. Nahezu alle Lehrpersonen würden das Verfahren in Zukunft einsetzen, wenn dies in elektronischer Form angeboten würde. Negativ wurde von einigen Lehrpersonen unter anderem die Ergebnisdarstellung, der erhöhte Zeitaufwand und die fehlende Anleitung zur Diskussion der Ergebnisse in der Klasse gesehen. Die Schülerinnen und Schüler schätzten an dem Verfahren, dass sie ihre Meinung äußern, somit positive oder negative Kritik üben konnten und damit halfen, den Unterricht zu verbessern. Ebenso wie die Lehrpersonen bemängelten die Schülerinnen und Schüler den benötigten Zeitaufwand.

### **5.3 Interklassenvergleich und Gruppenbildung**

Die im vorangegangenen Unterkapitel vorgestellten Ergebnisse werden nun eingehend analysiert. Als Ausgangspunkt dient der zentrale Zweck des Schülerfeedbacks, nämlich die Verbesserung der Unterrichtsqualität. Zunächst soll untersucht werden, ob und wie weit sich die Entwicklungen in den Skalen der Unterrichtsqualität zwischen den Klassen unterscheiden. Daraufhin werden aus Klassen, die einen ähnlichen Entwicklungsverlauf genommen haben, Gruppen gebildet. Mit Hilfe der weiteren Daten aus der Studie soll versucht werden innerhalb der Gruppen Muster zu identifizieren, welche die entsprechenden Entwicklungen erklären könnten. Dieses Vorgehen verfolgt das Ziel, differenzierte Hypothesen zu Wirkmechanismen des Schülerfeedbacks zu generieren.

#### **5.3.1 Interklassenvergleich, Skalen der Unterrichtsqualität**

Für den Interklassenvergleich der Unterrichtsqualitäts-Skalen werden für alle Klassen die Effektstärken  $d$  der Veränderungen zwischen Prä und Post ausgewiesen. Aus den Effektstärken lassen sich klare Rangfolgen bilden. Tabelle 26 zeigt die Effektstärken aggregiert über alle Klassen (letzte Zeile, siehe auch Abschnitt 5.2.3) und für jede einzelne Klasse als Zahl und als grüne (positive Veränderung) beziehungsweise rote Balken (negative Veränderung). Die Klassen sind nach ihrer durchschnittlichen Effektstärke über alle Skalen sortiert (letzte Spalte).

Tabelle 26: Interklassenvergleich Skalen der Unterrichtsqualität, sortiert nach der durchschnittlichen Effektstärke d

Unterrichtsqualität und Motivation aus Schülersicht (Effektstärken Prä:Post, grüne Balken entsprechen positiven und rote Balken negativen Effekten)													
Klasse	Verhältnis zur LP	Klarheit der LP	Inhaltliche Strukturiertheit	Formale kognitive Strukturiertheit	Angemessenheit Schwierigkeit	Angemessenheit Tempo	Diagnostische Kompetenz	Klassenmanagement	Bewertung der LP	Zeitnutzung	Interessanz	Motivierende Unterstützung durch LP	Durchschnittliche Effektstärke
10	0,88	0,70	0,49	0,48	0,35	0,44	0,89	0,28	0,72	0,04	0,40	1,11	0,57
8	0,42	0,42	0,30	0,16	0,35	0,37	0,90	0,10	0,08	0,30	1,01	0,68	0,42
7	0,33	0,60	0,36	0,49	0,25	0,25	0,39	0,08	0,79	0,18	0,09	0,67	0,37
4	0,51	0,16	0,55	0,53	-0,49	0,02	0,39	-0,20	0,33	0,26	0,12	0,48	0,22
2	0,50	-0,32	0,21	-0,01	0,25	0,44	0,15	-0,04	0,46	-0,10	0,47	0,10	0,18
6	-0,11	0,46	0,88	0,41	0,04	-0,18	0,05	0,06	0,13	-0,06	0,09	0,09	0,16
1	0,46	0,07	0,17	0,43	0,18	0,54	0,34	-0,89	0,19	-0,24	0,46	0,24	0,16
9	-0,49	0,28	0,44	0,69	-0,13	0,51	-0,46	0,02	-0,54	0,12	0,22	-0,34	0,03
5	0,04	-0,34	-0,03	-0,14	0,00	0,04	-0,06	-0,78	0,02	-0,12	0,24	0,59	-0,05
3	-0,46	-1,12	-0,05	-0,33	-0,58	0,19	-0,36	0,01	-0,99	0,04	-0,06	-0,37	-0,34
gesamt	0,14	0,07	0,27	0,19	-0,05	0,20	0,17	-0,05	0,04	0,02	0,21	0,21	0,12

Während Klasse 10 eine durchschnittliche Effektstärke von  $d = 0,57$  zeigt (nach Cohen ein mittlerer positiver Effekt), weist die Klasse 3 ein  $d$  von  $-0,34$  auf (mittlerer negativer Effekt). Insgesamt sieht man in sieben Klassen kleine bis mittlere positive Entwicklung, zwei Klassen zeigen insgesamt vernachlässigbare Effekte, und eine Klasse einen mittelgroßen negativen Effekt. Innerhalb der einzelnen Skalen (Spalten) treten teilweise deutlich größere Effekte beziehungsweise Effektunterschiede auf. Die Unterrichtsqualität hat sich demnach, sowohl zwischen den Klassen als auch zwischen den Skalen relativ heterogen entwickelt. Einen besonderen Status nimmt hierbei die schon erwähnte Klasse 3 ein. Sie weist die meisten und höchsten negativen Effekte auf und der Abstand zur nächstgelegenen Klasse ist mit durchschnittlichen 0,29 Punkten der weitaus größte.

Vergleicht man jedoch die Skalen innerhalb der Klassen und auch über die Klassen hinweg, so zeigt sich eine gewisse Regelmäßigkeit. Zwei Punkte fallen hierbei auf:

1. Die Tendenz der Skalen innerhalb der Klassen ist meist relativ homogen. Soll heißen, die meisten Skalen einer Klasse weisen, abgesehen von wenigen Ausnahmen, in die gleiche Richtung (positiv bzw. negativ). Als Besonderheit kann hier die Klasse 9 gesehen werden, die sowohl einzelne stark positive wie einzelne stark negative Effekte aufweist.
2. Zeigen Skalen einen positiven Effekt, dann sind es im Interklassenvergleich oft die gleichen (*Verhältnis zur Lehrperson, Klarheit der Lehrperson, Inhaltliche Strukturiertheit, Formal kognitive Strukturiertheit, Angemessenheit Tempo, Diagnostische Kompetenz, Bewertung der Lehrperson, Interessantheit und Motivierende Unterstützung*). Wohingegen die Skalen *Klassenmanagement, Angemessenheit Schwierigkeit* und *Zeitnutzung* nur in wenigen Klassen deutliche beziehungsweise meist nur kleine oder vernachlässigbare Effekte aufweisen. Dieses Muster lässt sich in Tabelle 26 als Eiszapfenprofil erkennen, geformt durch die grünen Balken.

Die Hypothese, dass ein Schülerfeedbackverfahren die Unterrichtsqualität positiv beeinflusst, scheint somit nicht für alle Skalen und auch nicht für jede Klasse zuzutreffen, beziehungsweise könnte der Einfluss auch negativ sein.

Die bisherige Betrachtungsweise hat die Ausgangswerte (Prä) nicht berücksichtigt. Diese spielen aber eine gewichtige Rolle. Ein Unterricht, der zu Prä bereits hoch bewertet wurde, besitzt ein geringeres Entwicklungspotential (ausgedrückt in Cohens d) als ein Unterricht der zu Prä niedriger bewertet wurde. Aus diesem Grund werden in Tabelle 27 die Ausgangswerte der Skalen aller Klassen zu Prä aufgeführt. Die Sortierreihenfolge ist die selbe wie in Tabelle 26; als Referenz wurden die durchschnittlichen Effektstärken in der zweiten Spalte nochmals aufgeführt. Die letzte Spalte stellt die Durchschnittswerte aller Skalen zu Prä in den einzelnen Klassen dar. Je dunkler die Felder sind, desto höher sind die jeweiligen Werte.



Tabelle 27: Unterrichtsqualität, Vergleich der durchschnittlichen Effektstärke mit den Werten zur Präerhebung

Klasse	Durchschnittliche Effektstärke der Unterrichtsqualität (Prä-Post)	Messwerte Skalen der Unterrichtsqualität zu Prä (4-Stufige Likertskala, je dunkler, desto höher)												
		Verhältnis zur LP	Klarheit der LP	Inhaltliche Strukturiertheit	Formale kognitive Strukturiertheit	Angemessenheit Schwierigkeit	Angemessenheit Tempo	Diagnostische Kompetenz	Klassenmanagement	Bewertung der LP	Zeitnutzung	Interessanz	Motivierende Unterstützung durch LP	Mittelwerte der Skalen
10	0.57	2.37	2.84	2.29	2.55	2.82	2.97	2.43	2.53	2.92	2.90	2.48	2.78	2.66
8	0.42	2.61	3.29	2.64	2.89	2.73	3.17	2.56	3.51	3.33	3.11	2.69	3.05	2.96
7	0.37	3.05	3.21	2.62	2.99	3.11	2.86	3.00	3.31	3.09	3.20	3.02	2.86	3.03
4	0.22	2.80	3.64	2.76	3.17	2.84	3.44	2.81	3.40	3.62	3.12	3.08	3.25	3.16
2	0.18	3.12	3.75	2.76	3.23	3.23	3.18	3.22	3.36	3.35	3.57	2.77	3.43	3.25
6	0.155	2.97	3.39	2.60	3.04	3.24	3.20	2.99	3.42	3.23	3.32	2.98	3.19	3.13
1	0.163	2.66	3.56	2.79	2.92	2.68	2.84	2.90	3.67	3.60	2.93	2.66	3.07	3.02
9	0.03	2.98	3.19	2.34	2.64	3.19	2.97	2.95	3.07	3.25	3.31	2.69	3.12	2.97
5	-0.05	2.79	3.51	2.91	3.11	3.18	3.27	2.72	3.58	3.42	3.30	2.80	2.97	3.13
3	-0.34	2.49	2.94	2.11	2.38	3.02	1.82	2.27	1.53	2.49	2.78	2.21	2.67	2.39

Bezüglich der Ausgangsmittelwerte (letzte Spalte), lässt sich eine Tendenz zur Mitte erkennen. Die Klassen mit den höchsten Mittelwerten liegen bezüglich Rangfolge der Effekte (2. Spalte) im Mittelfeld. Werden beide Informationen miteinander kombiniert, lassen sich fünf Gruppen identifizieren. Allerdings sind die Klassen 1 und 5 nicht eindeutig zuzuordnen. Geht man ausschließlich von den Präwerten aus, würden die Klassen 1 und 9 eine Gruppe bilden und Gruppe 5 würde den Klassen 4, 2 und 6 zugeordnet; ginge man allein von den Effekten aus, würden die Klassen 5 und 9 zu einer Gruppe zusammengefasst und Gruppe 1 wäre Teil der Gruppe 2, 4 und 6. In der folgenden Darstellung werden deshalb die Gruppen 1 und 5 in den jeweiligen Gruppen mit eckiger Klammer dargestellt:

1. Eine Klasse, mit einem unterdurchschnittlichen Wert zu Prä und einem vergleichsweise großen positiven Effekt (Klasse 10).
2. Klassen mit einem im Quervergleich unauffälligen Wert von ungefähr 3 zu Prä und einem mittelgroßen positiven Effekt. (Klassen 7 und 8)
3. Klassen mit einem leicht überdurchschnittlichen Wert von 3 bis 3.25 zu Prä und einem kleinen positiven Effekt. (Klassen [1], [5], 2, 4 und 6)

4. Klassen mit einem im Quervergleich unauffälligen Wert von ungefähr 3 zu Prä und keinem messbaren Effekt. (Klassen [1], [5] und 9)
5. Eine Klasse, mit einem stark unterdurchschnittlichen Wert zu Prä und einem mittelgroßen negativen Effekt. (Klasse 3)

Tabelle 28: Clusteranalyse

Clusterzugehörigkeit	
Fall	5-Cluster
1	4
2	3
3	5
4	3
5	4
6	3
7	2
8	2
9	4
10	1

Es wurde mittels Clusteranalyse versucht, einerseits die unscharfen Gruppenzuordnungen mathematisch zu lösen und andererseits die eindeutigen zu validieren. Hierfür wurden die klassenweisen arithmetischen Mittelwerte der Effektstärken (Tab. 27, Spalte 2) und die klassenweisen arithmetischen Mittelwerte zu Prä z-standardisiert und basierend auf der quadrierten euklidischen Distanz der Einzelwerte, Cluster gebildet. Dies ergab eine Zusammenfassung der Klassen 1, 5 und 9 zu einer Gruppe. Das generelle deskriptive Muster wurde bestätigt (siehe Abb. 10 und Tab. 28).

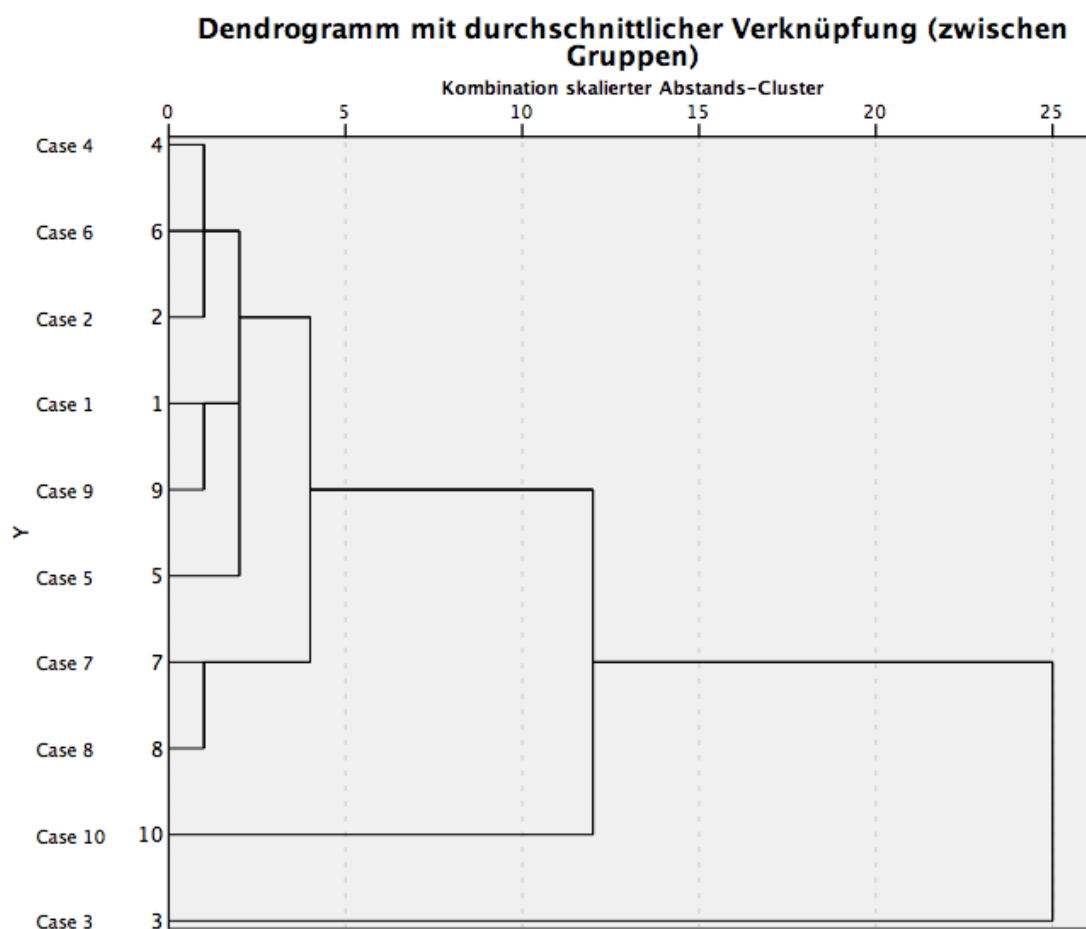


Abbildung 10: Dendrogramm, Clusteranalyse Effektstärken und Präerhebung der Unterrichtsqualität

Folgende Gruppen wurden demnach gebildet:

1. Eine Klasse mit einem unterdurchschnittlichen Wert zu Prä und einem vergleichsweise hohen positiven Effekt: **Klasse 10**
2. Klassen mit einem durchschnittlichen Wert zu Prä und einem vergleichsweise überdurchschnittlichen Effekt, bestehend aus: **Klasse 8, Klasse 7**
3. Klassen mit einem hohen Wert zu Prä und einem vergleichsweise durchschnittlichen positiven Effekt, bestehend aus: **Klasse 4, Klasse 2, Klasse 6**
4. Klassen mit einem durchschnittlichen Wert zu Prä und einem vergleichsweise unterdurchschnittlichen Effekt, bestehend aus: **Klasse 1, Klasse 9, Klasse 5**
5. Eine Klasse, mit einem unterdurchschnittlichen Wert zu Prä und einem vergleichsweise hohen negativen Effekt: **Klasse 3**

Diese Gruppen dienen als Startpunkt für die weitere Datenanalyse zur Identifikation von Mustern im Unterrichtsprozess. Ergäbe sich im weiteren Verlauf der Datenanalyse ein besserer Fit durch eine Veränderung der Gruppen, würden diese entsprechend angepasst.

## 5.4 Intergruppenvergleich

### 5.4.1 Intergruppenvergleich, Verlauf Schüler-Kurzfeedbacks

In diesem Abschnitt wird der Frage nachgegangen, inwiefern sich Muster, die sich durch die Prä-/Posterhebung der Unterrichtsqualität ergeben haben, auch in den Verläufen der Schüler-Kurzfeedbacks widerspiegeln. In der Auswertungsmethodik und -interpretation gleicht die Vorgehensweise der Auswertung über alle Schulklassen (siehe Abschnitt 5.2.5).

Betrachtet man Tabelle 29 gesamthaft, fällt zunächst auf, dass mit Ausnahme des Diagramms in Spalte 1, Zeile 3 keine signifikant negative Steigungen festzustellen sind. Die Mehrzahl der Diagramme weist positive Steigungswerte auf, die wiederum in acht Fällen das Signifikanzniveau von mindestens 5 % erreichen. Außerdem zeigt sich, dass die, bereits in Abschnitt 5.2.5 dargestellte signifikante Veränderung der Skalen *Disziplin* und *Experimentieren als theoriegeleitetes Handeln*, nicht aufgrund einer einzelnen starken Veränderung einer Gruppe oder Klasse stattgefunden hat, sondern, dass sich diese signifikante Veränderung in über der Hälfte der Gruppen wiederfindet. Schülerinnen und Schüler aus relativ vielen unterschiedlichen Klassen haben demnach eine Veränderung dieser Skalen wahrgenommen.

Im Folgenden werden die einzelnen Skalen (vgl. Spalten der Tabelle 29) im Gruppenvergleich genauer betrachtet (folgende Abschnitte 5.4.1.1 – 5.4.1.5), um daraufhin die einzelnen Gruppen (vgl. Zeilen der Tabelle 29) in ihrem Muster zu beschreiben und mit den bereits gefunden Mustern aus der Prä-/Posterhebung zu vergleichen (Abschnitt 5.4.1.6).

Tabelle 29: Verläufe der Kurzfeedbacks nach Gruppen getrennt (schwarze durchgehende Linie): Regressionsgerade mit Steigungsgleichung

	Kognitive Passung (nicht überfordert)	Kognitive Passung (nicht unterfordert)	Kognitiv aktivierende und strukturierende Elemente	Disziplin	Experimentieren als Theoriegeleitetes Handeln
Gruppe 1	<p>Signifikanzniveau p der Steigung = 0,036*</p>	<p>Signifikanzniveau p der Steigung = 0,789</p>	<p>Signifikanzniveau p der Steigung = 0,027*</p>	<p>Signifikanzniveau p der Steigung = 0,000***</p>	<p>Signifikanzniveau p der Steigung = 0,007*** (gemessene Werte nur MZ 1 und MZ 3)</p>
Gruppe 2	<p>Signifikanzniveau p der Steigung = 0,326</p>	<p>Signifikanzniveau p der Steigung = 0,053</p>	<p>Signifikanzniveau p der Steigung = 0,303</p>	<p>Signifikanzniveau p der Steigung = 0,026*</p>	<p>Signifikanzniveau p der Steigung = 0,930</p>
Gruppe 3	<p>Signifikanzniveau p der Steigung = 0,000***</p>	<p>Signifikanzniveau p der Steigung = 0,795</p>	<p>Signifikanzniveau p der Steigung = 0,421</p>	<p>Signifikanzniveau p der Steigung = 0,516</p>	<p>Signifikanzniveau p der Steigung = 0,013*</p>
Gruppe 4	<p>Signifikanzniveau p der Steigung = 0,425</p>	<p>Signifikanzniveau p der Steigung = 0,461</p>	<p>Signifikanzniveau p der Steigung = 0,641</p>	<p>Signifikanzniveau p der Steigung = 0,062</p>	<p>Signifikanzniveau p der Steigung = 0,495</p>
Gruppe 5	<p>Signifikanzniveau p der Steigung = 0,398</p>	<p>Signifikanzniveau p der Steigung = 0,057</p>	<p>Signifikanzniveau p der Steigung = 0,115</p>	<p>Signifikanzniveau p der Steigung = 0,001**</p>	<p>Signifikanzniveau p der Steigung = 0,000**</p>

#### **5.4.1.1 Kognitive Passung (nicht überfordert)**

Gruppe 1 weist in der Skala *Kognitive Passung (nicht überfordert)* als einzige eine signifikante positive Entwicklung auf, allerdings ausgehend von einem vergleichsweise niedrigen Wert zu MZ 1 (2,68). Gruppen 2, 4 und 5 zeigen keine signifikante Veränderung, allerdings fällt bei Gruppe 5 auf, dass der Ausgangswert, ähnlich wie in Gruppe 1, relativ tief liegt (2,98). Die Gruppe 3 verläuft negativ, allerdings findet dieser Trend auf so hohem Niveau statt, dass sich selbst der Wert zu MZ 4 im Vergleich zu den anderen Gruppen im Mittelfeld befindet. Die größte Schwankung der Messwerte ist in den Gruppen 1 und 5 festzustellen.

Die Gruppen 2 und 4 lassen sich also bezüglich des Verlaufs und der Niveaus der Messwerte zusammenfassen. Aufgrund der fallenden Tendenz kann man Gruppe 3 abgrenzen. Ebenso die Gruppen 1 und 5, welche in diesem Vergleich eindeutige Unterscheidungsmerkmale zeigen. Bei beiden sind die Werte zu MZ 1 wesentlich tiefer angesiedelt, allerdings verzeichnet Gruppe 1 als einzige eine signifikant positive Entwicklung über den Messzeitraum.

Das Muster der Prä-/Posterhebung der Unterrichtsqualität lässt sich also teilweise in den Verläufen dieser Skala wiederfinden (vor allem in den Gruppen 1 und 4; bezogen auf das Ausgangsniveau auch Gruppe 2 und 5). Der negative Trend der Gruppe 5 findet sich hier jedoch nicht wieder. Umgekehrt verhält es sich auf den ersten Blick mit der Gruppe 3, deren negativer Trend nicht im Einklang mit der Prä-/Posterhebung steht, aufgrund des hohen Niveaus fällt dieser Widerspruch jedoch nicht stark ins Gewicht.

#### **5.4.1.2 Kognitive Passung (nicht unterfordert)**

Vergleicht man die Verläufe der *Kognitiven Passung (nicht unterfordert)* der einzelnen Gruppen, fällt auf, dass sich in den ersten vier Gruppen die Werte relativ konstant auf hohem Niveau bewegen. Die Achsenabschnitte liegen hier über dem Wert von 3,0 und unterschreiten diesen Wert nie. Weder ist eine signifikante Steigung zu beobachten, noch starke Ausschläge zu den jeweiligen Messzeitpunkten. Insgesamt spiegeln die vier Gruppen den Verlauf über alle Klassen dieser Skala aus Abschnitt 5.2.5 wider.

Gruppe 5 weicht von diesem Gesamtbild ab. Die Werte dieser Gruppe liegen deutlich tiefer. Außerdem fällt auf, dass sie zu den jeweiligen Messzeitpunkten stärker variieren. Die Steigung ist knapp nicht signifikant.

Es lässt sich also festhalten, dass bei qualitativer Betrachtungsweise einzig Gruppe 5 gegenüber den anderen Gruppen auffällt, nämlich durch ein relativ tiefes Niveau dieser Skala. Eine deutliche negative Entwicklung wie sie bei der Prä-/Posterhebung der Unterrichtsqualität festgestellt wurde (siehe Unterkapitel 5.3) ist hier jedoch nicht erkennbar.

#### **5.4.1.3 Kognitiv aktivierende und strukturierende Elemente**

Die Skala *Kognitiv aktivierende und strukturierende Elemente* zeigt bei den Gruppen 2 bis 4 Ähnlichkeiten zu den zuvor analysierten Skalen. Die Werte liegen auf relativ hohem Niveau, es sind keine positiven oder negativen Entwicklung festzustellen und die Werte zeigen eine relativ hohe Konstanz über die vier Messzeitpunkte.

Gruppe 1 unterscheidet sich von dem beschriebenen Muster und verhält sich ähnlich wie in der Skala *Kognitive Passung (nicht überfordert)*. Der Ausgangswert ist unterdurchschnittlich, es folgt aber eine signifikant positive Entwicklung. Auch Gruppe 5 weist einen niedrigen Ausgangswert auf, die leicht positive Entwicklungstendenz ist aber nicht signifikant. Diese beiden Gruppen zeigen auch hier die größten Schwankungen zwischen den vier Messzeitpunkten.

Der Trend, dass sich die Gruppen 2 bis 4 sehr ähnlich sind und sich die Gruppen 1 und 5 am deutlichsten von diesen unterscheiden lassen, setzt sich in dieser Skala fort. Während jedoch Gruppe 1 erneut ein ähnliches Muster aufweist wie im Prä-/Post-Vergleich, finden die Werte der Gruppe 5 abermals nur bezüglich des Ausgangsniveaus der Prä-/Posterhebung eine Entsprechung. Eine signifikant negative Entwicklung der Skala ist nicht festzustellen.

#### **5.4.1.4 Disziplin**

Die Verläufe der Skala *Disziplin* zeigen, wie schon erwähnt, in vier Gruppen eine signifikante (bzw. in einem Fall nahezu signifikante) positive Entwicklung. Einzig Gruppe 3 weist keine positive Entwicklung auf, befindet sich aber insgesamt auf sehr hohem Niveau. Die eindeutig stärkste positive Entwicklung erkennt man auch hier in Gruppe 1, gefolgt von Gruppe 5. Beide Gruppen starten aber (erneut) bei vergleichsweise deutlich tieferen Ausgangswerten.

Auch in dieser Skala ist der bereits vorher beschriebene Trend (die relativ ähnlichen Verläufe der Gruppen 2 bis 4), wenngleich weniger eindeutig, sichtbar. Ebenfalls fällt auf, dass Gruppe 5 eine signifikant positive Entwicklung nimmt. Erneut kann also die negative Entwicklung aus dem Prä-/Postvergleich in den Kurzfeedback-Verläufen nicht bestätigt werden. Im Gegensatz dazu bildet die Gruppe 1 wiederholt den Trend der Prä-/Posterhebungen im Schüler-Kurzfeedback ab.

#### **5.4.1.5 Experimentieren als theoriegeleitetes Handeln**

Die Gruppen 1, 3 und 5 zeigen in dieser Skala signifikant positive Entwicklungen. Wobei zu beachten ist, dass in Gruppe 1 nur zu MZ 1 und MZ 3 Experimente durchgeführt wurden und somit die Aussagekraft des Verlaufs eingeschränkt ist. Gruppe 3 weist im Inter-Skalenvergleich zwei Auffälligkeiten auf: Erstens besitzt diese im Vergleich zu den bisher vorgestellten Skalen den niedrigsten Ausgangswert (MZ 1 ist kaum höher als in Gruppe 5) und zweitens ist dies die einzige Skala, in der diese Gruppe eine signifikant positive Entwicklung aufzeigt.

Die Messwerte der Gruppe 2 und 4 verhalten sich auf vergleichsweise hohem Niveau relativ unauffällig und konstant und zeigen somit Ähnlichkeiten zu den Messungen der vorigen Skalen. Ebenso Gruppe 5, die erneut auf relativ tiefem Niveau beginnt, sich dann aber signifikant positiv entwickelt und somit erneut gegen den Trend der Prä-/Posterhebung verläuft.

#### **5.4.1.6 Zusammenfassung des Intergruppenvergleichs der Kurzfeedback-Verläufe**

Betrachtet man die Verläufe der einzelnen Gruppen (Zeilen der Tabelle 29) über alle Skalen, so erkennt man, dass die Gruppe 1 in vier von fünf Skalen auf vergleichsweise unterdurchschnittlichem Niveau beginnt, daraufhin aber signifikant positive Entwicklungen (jeweils die höchsten aller Gruppen) aufweist. Die Verläufe der Kurzfeedbacks des Gruppen 1 ähneln also in starkem Maße der Entwicklung der gemessenen Ausgangswerte und Effekte in der Prä-/Posterhebung der Unterrichtsqualität.

In der Gesamtschau der Gruppen 2 bis 4 scheint dies auf den ersten Blick nicht so eindeutig zu sein. Bei genauerem Hinsehen wird aber auch hier ein sehr ähnlicher Verlauf zur Unterrichtsqualität deutlich. So zeigt Gruppe 3 in vier von fünf Fällen den höchsten Wert zu MZ 1. Doch auch die Skala *Experimentieren als theoriegeleitetes Handeln*, die zu MZ 1 einen unterdurchschnittlichen Wert in Gruppe 3 aufweist, steigt zu MZ 2 stark an. Insgesamt wurde also ein überdurchschnittliches Niveau nochmals verbessert. Selbst die höchstsignifikant fallende Tendenz in der Skala *Kognitive Passung (nicht überfordert)* dient nur als schwaches Gegenargument, da sich der Wert zu MZ 4, im Vergleich zu den anderen Gruppen immer noch auf mindestens durchschnittlichen Niveau bewegt.

Die Gruppen 2 und 4 zeigen in allen Skalen relativ hohe Werte jedoch nur einmal den Höchstwert aller Gruppen. Gruppe 2 unterscheidet sich von Gruppe 4 dahingehend, dass zumindest in einer Skala eine positive Entwicklung festzustellen ist (*Disziplin*), in Gruppe 4 ist dies nie der Fall. Auch in diesen beiden Gruppen lassen sich also Ähnlichkeiten zum Verlauf der Unterrichtsqualität (Prä-Post) finden, denn beide Skalen zeigten bei der Präerhebung durchschnittliche Werte, die Werte von Gruppe 2 stiegen jedoch zur Posterhebung stärker an.

Einzig Gruppe 5 weist in ihrer Entwicklung ein widersprüchliches Bild zur erhobenen Unterrichtsqualität (Prä-Post) auf. Zwar liegen die Werte auch während der Kurzfeedbacks zu MZ 1 auf unterdurchschnittlichem Niveau, im Gegensatz zur Entwicklung der Unterrichtsqualität zeigt hier keine Skala einen negativen Verlauf, in drei Skalen kommt es sogar zu positiven Entwicklungen.

#### **5.4.1.7 Intergruppenvergleich der offenen Fragen im Schülerkurzfeedback**

Im Laufe der Intervention hatten die Schülerinnen und Schüler die Möglichkeit, im Kurzfeedback-Fragebogen auf zwei offene Fragen zu antworten. Während sich die erste auf die Inhalte des Unterrichts bezog und ob die Schülerinnen und Schüler diesbezüglich Verständnisprobleme

wahrnahmen (Item OFR01: *Folgende Themen des heutigen Unterrichts sollten nochmals angesprochen werden, damit ich sie verstehe*), war die zweite Frage allgemeiner formuliert (Item OFR02: *Was ich unbedingt noch zu der heutigen Schulstunde sagen will*). Man kann davon ausgehen, dass die Schülerinnen und Schüler einer Gruppe, welche hohe Werte bei der Unterrichtsqualität aufweist, seltener eine Wiederholung eines Themas fordern (OFR01), beziehungsweise weniger negative und mehr positive Kommentare bezüglich der gehaltenen Unterrichtsstunde abgeben (OFR02).

Um dies zu untersuchen, wurden die Antworten der jeweiligen Items kategorisiert und gezählt. Bei Item OFR01 wurde zwischen sachlichem (z. B. „Formeln wiederholen“) und unsachlichem (z. B. „Ich will Eis“) Kommentar unterschieden, Item OFR02 differenzierte sich in positiv (z. B. „gute Atmosphäre“), negativ (z. B. „Allgemein war es zu laut“), unsachlich (s. o.) und ambivalent (z. B. „Besser als letzte Stunde, aber immer noch laut“) unterschieden. Die Anzahl der Nennungen wurde durch die Anzahl der Schülerinnen und Schüler in der jeweiligen Gruppe dividiert um vergleichbare relative Häufigkeiten zu erhalten.

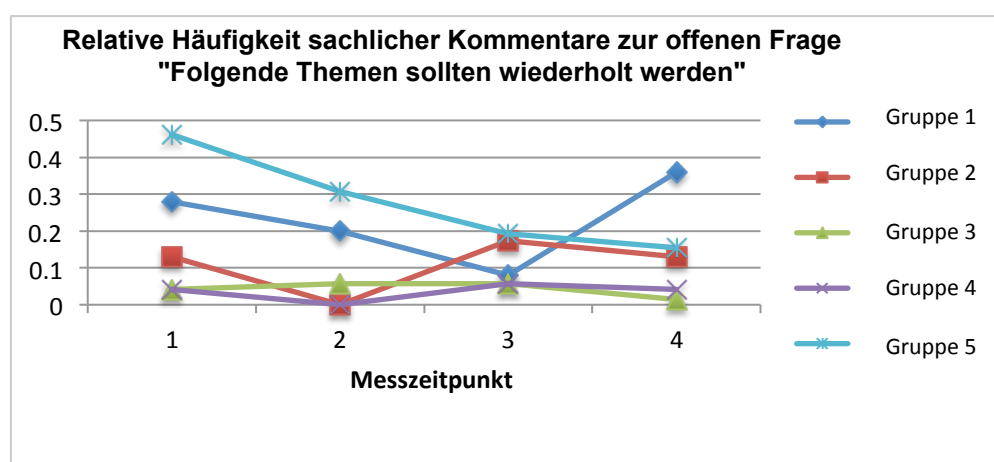


Abbildung 11: Offene Antworten im Schüler-Kurzfeedback: Wiederholung (sachlich)

Abb. 11 zeigt den Verlauf der sachlichen Kommentare des Items OFR01 über die vier Messzeitpunkte der fünf Gruppen. Zu Beginn liegen die Werte zwischen den Gruppen relativ weit auseinander, um sich im Verlauf der Feedbackphase auf niedrigem Niveau tendenziell anzugleichen. Diese Tendenz wird allerdings durch den Anstieg der Gruppe 1 zu Messzeitpunkt 4 abgeschwächt. Das Muster aus der Prä-/Posterhebung lässt sich vor allem zu Messzeitpunkt 1 wiedererkennen. Nahezu jeder zweite (Gruppe 5) beziehungsweise jeder dritte (Gruppe 1) Schüler forderte die Lehrperson auf, Teile des behandelten Stoffs in der nächsten Stunde zu wiederholen. Bei den Schülerinnen und Schülern der Gruppe 2 bis 4 war dies maximal jeder siebte. Das schon erwähnte Angleichen der Werte im Verlauf des Untersuchungszeitraums führt zu widersprüchlicher Tendenz der Gruppe 5 (Abnahme der relativen Häufigkeit der Nennungen) im Vergleich zur Entwicklung der Unterrichtsqualität (negativer Effekt im Prä-



/Postvergleich). Ähnlich widersprüchlich ist der Anstieg der Gruppe 1 zu Messzeitpunkt 4, der jedoch durch das Interview mit der Lehrperson geklärt werden konnte (Einstieg in ein neues Thema).

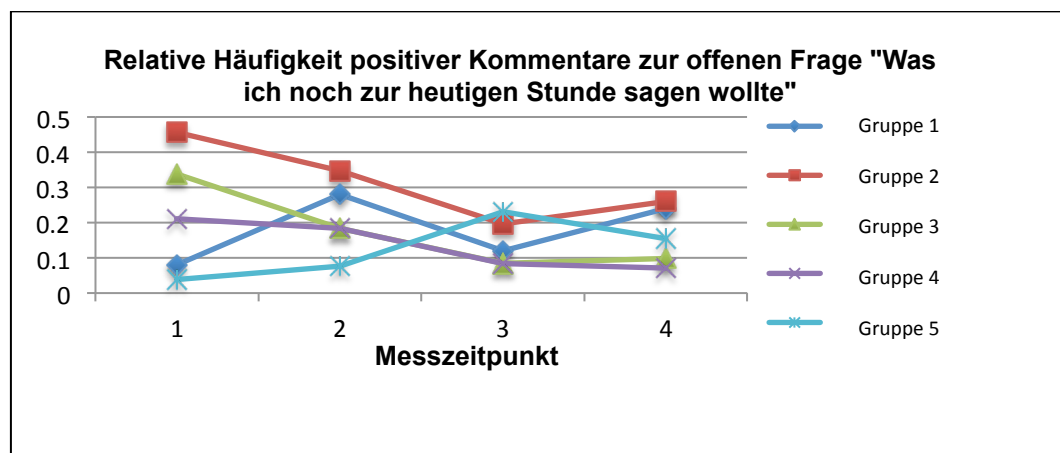


Abbildung 12: Offene Antworten im Schüler-Kurzfeedback: Allgemein (positiv)

Ein ähnliches Bild ergibt sich, wenn man den Verlauf der positiven Kommentare des Items OFR02 betrachtet. So kommentierten erwartungsgemäß die Schülerinnen und Schüler der Gruppe 2 bis 4 zu MZ 1 häufiger positiv (nahezu jeder zweite, bzw. mehr als jeder fünfte, gegenüber weniger als einem Zehntel der Schülerinnen und Schüler in Gruppe 1 und 5). Allerdings fallen beim weiteren Verlauf kaum erwartungskonforme Muster auf. Die Gruppen 1 bis 3, deren positive Effektstärken aufgrund des Prä-/Postvergleich der Unterrichtsqualität überdurchschnittlich ausgeprägt waren, zeigen eher eine sinkende beziehungsweise konstante Tendenz, während Gruppe 5 zu MZ 4 mehr positive Kommentare aufweist als zu MZ 1.

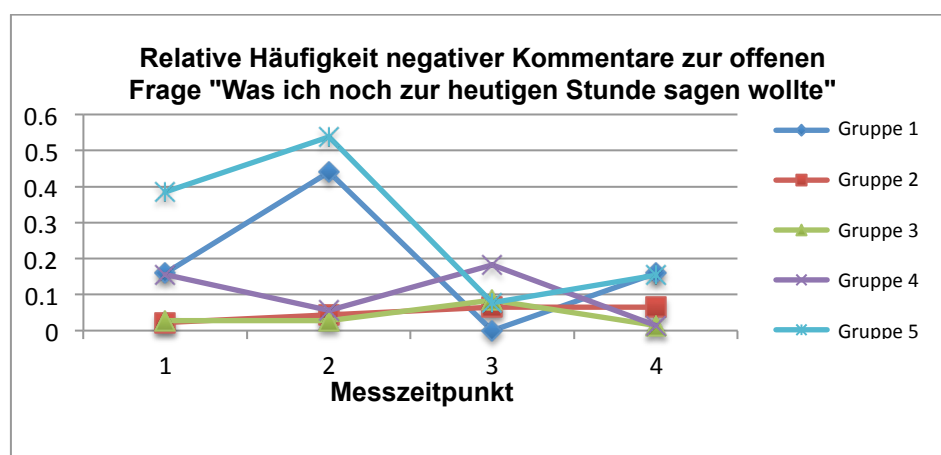


Abbildung 13: Offene Antworten im Schüler-Kurzfeedback: Allgemein (negativ)

Bei der Betrachtung der Häufigkeit negativer Kommentare (Abb.13) ergibt sich ein nahezu erwartungsgemäßes Bild zu MZ 1. Die Schülerinnen und Schüler der Gruppen 1 und 5 formulier-

ten zu MZ 1 die meisten negativen Kommentare, wenngleich Gruppe 4 einen ähnlich hohen Wert aufweist wie Gruppe 1. Der weitere Verlauf der beiden Gruppen variiert zunächst stark, nimmt dann aber zum Ende ähnliche Werte wie die anderen Gruppen an.

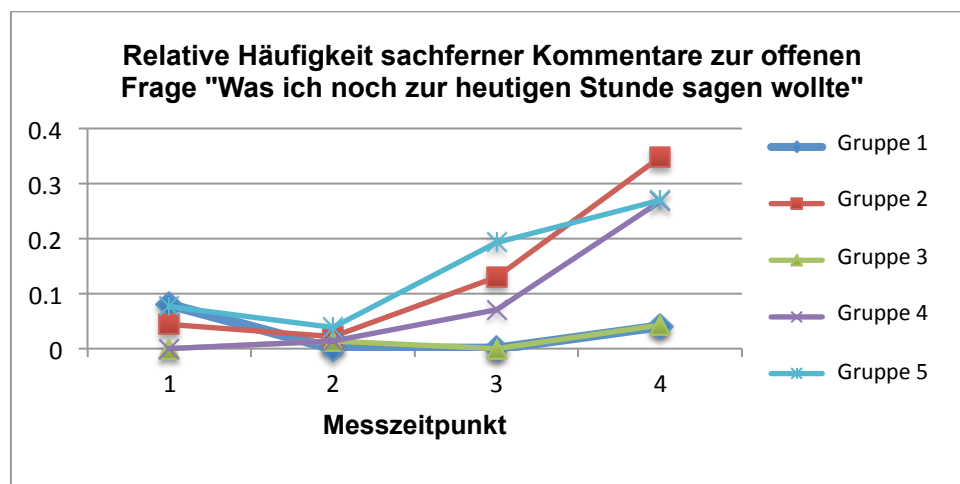


Abbildung 14: Offene Antworten im Schüler-Kurzfeedback: Wiederholung (sachfern)

Insgesamt zeigt sich in allen Kategorien eine abnehmende Anzahl sachlicher Kommentare innerhalb der beiden offenen Antwortformate. Dies könnte mit einer, über den Messzeitraum sich einstellenden Antwortmüdigkeit der Schülerinnen und Schüler zu erklären sein. Außerdem ist erkennbar, dass in einigen Gruppen (genauer in einigen Klassen) die sachfernen Kommentare in beiden offenen Items über den Messzeitraum zunahmen (vor allem OFRA02, siehe Abb. 14). Hier fallen hauptsächlich die Gruppen 2, 4 und 5 auf. Bei genauerer Analyse der Kommentare fällt jedoch auf, dass diese sich zwischen den einzelnen Gruppen unterscheiden. Während in den Gruppen 4 und 2 die Schülerinnen und Schüler häufig nicht änderbare Rahmenbedingungen kommentierten („es war zu heiß“) oder versuchten humorvolle Kommentare zu verfassen beziehungsweise Komplimente zu machen (z. B. „Herrn x stehen rote Pullis echt gut“), waren die Kommentare in Gruppe 5 häufig destruktiv oder in einem „Insider-Sprachcode“ verfasst (z. B. „Poa ey, immer Mathematik altaa...20-30 tausend Einwohner desch wie Yoga! Puuh!“). Dies kann als demonstratives Desinteresse am Unterricht und/oder am Feedbackverfahren interpretiert werden. Dieses Ergebnis scheint mit der Klage einiger Schülerinnen und Schüler dieser Gruppe im Einklang zu stehen, die bei manchen Mitschüler/-innen die Ernsthaftigkeit gegenüber dem Verfahren vermissten (siehe Abschnitt 5.4.2).

Der Anstieg der unsachlichen Kommentare in manchen Gruppen (bzw. Klassen) zum Ende des Messzeitraumes wird hauptsächlich von zwei Sachverhalten begünstigt: Zunächst beginnen einige wenige Schülerinnen und Schüler in einer Schulstunde unsachliche Kommentare schreiben (hier MZ 3), woraufhin die Lehrperson diese in der darauffolgenden Stunde zeigt. Dies

scheint weitere Schülerinnen und Schüler zu ermuntern, ebenfalls unsachliche Kommentare zu verfassen. So beschrieb es die Lehrperson der Klasse 7 (Gruppe 2) im Interview:

*I Wie haben Sie denn auf diese, zu Messzeitpunkt 3 gab es so die ersten Quatschkommentare, in Anführungszeichen, ehm wie haben Sie jetzt darauf reagiert? Haben Sie auch alle gezeigt?*

*B Ich habe sie vereinzelt, vorgelesen, ich eh, also das war natürlich einfach so ne Sache, ich habe mir im Nachhinein gedacht, eigentlich hätte ich sie gar nicht erwähnen sollen, weil das hat die Schüler dann angestachelt noch mehr von diesen eh Quatschkommentaren eh zu schreiben und sich zu produzieren, ehm ja also, ich hab, das Feedback ist halt auch eh, in einer sechsten Stunde dann zum Beispiel geschrieben worden und da waren halt die Schüler dann schon auch sehr in Wochenendfeeling, das muss man wirklich auch eh, sehen, und ja, vielleicht war es ihnen auch zu blöd, dass sie nichts gewusst haben zu schreiben und dann besser irgendein Blödsinn als gar nichts und eh, ich muss dazu auch sagen, sie sind auch nach wie vor von der Qualität der Witze überzeugt. Und*

*I (<lacht>)*

*B immer noch sehr stolz darauf. Ich meine, dass zumindest ein paar wirklich gelungen waren.*

*I (<lacht>)*

*B Also, es ist halt eine neunte Klasse.*

Aus dem Subtext dieses Interviewabschnitts lässt sich herauslesen, dass die Lehrperson die unsachlichen Kommentare zwar nicht zielführend fand, dafür aber Verständnis aufbrachte und sich und seine Arbeit nicht herabgewürdigt sah. Als Gegenbeispiel soll hier der diesbezügliche Interviewausschnitt aus Gruppe 5 (Klasse 3) dienen:

*B Diese Kommentarfunktion, die ist ja rege genutzt worden auch für Quatsch machen und ehm ja, das konnte ich leider nicht alles ganz ernst nehmen, was da stand. Aber ich habs natürlich gezeigt, die haben sich auch sehr gefreut, dass dann ihr Kommentar eh da vorkam und dann sagte einer, nein, meins fehlt, meines fehlt, was ich geschrieben habe und ehm jetzt stand es dann aber doch da. Es hat nicht gefehlt. Es war da, stand nur woanders.*

*I Würden Sie denn, wenn Sie jetzt angenommen eh in Zukunft diese Quatschkommentare immer noch zeigen, wenn Sie das jetzt nochmal anwenden würden, dieses Feedback.*

**B** Ja, ich würde das zeigen, ich habe es auch jedes Mal gezeigt, hat ja nicht groß nachgelassen der Quatsch, ehm ich würde es deswegen zeigen, damit, damit jeder in der Klasse sieht, auf welchem Niveau wir hier gerade uns bewegen, und wie wenig ernst wir das nehmen denn das ist für mich natürlich auch ein Mittel zu zeigen, ja es liegt jetzt nicht nur, es liegt jetzt nicht am an meinem Unterricht also, nicht ausschließlich, wird ja gern so gesehen von Schülern, dass eh wurde mir auch gesagt von einigen, sie müssen halt interessanter unterrichten. Dann hören wir auch zu.

**I** Ja.

**B** Ehm aber genau diese Kommentare sind natürlich ein Hinweis darauf, dass die Ernsthaftigkeit jetzt nicht bei allen gegeben ist und da war mir das eigentlich sogar willkommen.

**I** Ja, ja.

**B** Ja und da kann man auch immer sagen, ja wenn solche Kommentare da darauf geschrieben werden dann habe ich keine Fragen mehr, keine weiteren Fragen.

**I** Ja, ja.

**B** Und deswegen fand ich das jetzt nicht verkehrt und hab das auch gezeigt und haben sich alle auch gefreut dann daran was da dummes geschrieben wurde, ist ja auch mal lustig aber eh ja, hilft jetzt inhaltlich natürlich nicht.

Zum einen sieht man an diesem Ausschnitt, dass die Lehrperson das eigene Verhalten bezüglich der Vorstellung der unsachlichen Kommentare nicht anzweifelte und zum anderen fiel es der Lehrperson deutlich schwerer, die unsachlichen Kommentare als normales Schülerverhalten zu interpretieren, was wahrscheinlich an der schon angesprochenen "Qualität" der Kommentare lag.

Insgesamt lassen sich auch im offenen Antwortformat teilweise erwartungskonforme Muster erkennen, die zu den Mustern passen, welche im Prä-/Postvergleich sowie in den Kurzfeedbackverläufen gefunden wurden. Dabei sind es erneut die Gruppen 1 und 5, die sich relativ stark von den Gruppen 2 bis 4 unterscheiden.

Fasst man den Intergruppenvergleich der offenen und geschlossenen Antwortformate der Schüler-Kurzfeedbacks zusammen, lässt sich festhalten, dass sich das beobachtete Muster der Unterrichtsqualitäts-Entwicklung (Prä-Post) zum größten Teil bei den Schüler-Kurzfeedbacks wiederfinden lässt (insbesondere hinsichtlich Ausgangswerten, Entwicklungstendenzen und Häufigkeiten von Antwortkategorien bei offenen Items). Eine auffällige Ausnahme bildet die Gruppe 5, die zwar zwischen Prä-/Post-Vergleich und den Kurzfeedbacks in vielen Skalen vergleichba-

re Ausgangsniveaus aufwies, jedoch hinsichtlich der Entwicklung (Prä-/Postvergleich klar negative Effekte; Kurzfeedbackverläufe in der Tendenz positiv) Widersprüchlichkeiten zeigte.

#### 5.4.2 Intergruppenvergleich, Einschätzung des Schüler-Kurzfeedback-Verfahrens durch die Schülerinnen und Schüler

Tabelle 30 listet die Anzahl der Nennungen des Items SWSF21 auf (*Welche Veränderungen im Unterricht sind dir besonders aufgefallen?* ausschließlich berücksichtigt wenn SWSF02 mit mindestens "stimme eher zu" beantwortet wurde). Um die Werte der Gruppen vergleichbar zu machen, wird die relative Anzahl der Nennungen (pro Schüler/in) aufgeführt. Insgesamt fällt auf, dass die Gruppen 1 und 5 überdurchschnittlich viele Nennungen sowohl insgesamt als auch bezüglich der Verbesserung der Disziplin aufweisen.

Tabelle 30: Clustervergleich Anzahl der Nennungen auf die offene Frage des Items V2SWSF21

Relative Anzahl Nennungen offene Frage: <i>Welche Veränderungen im Unterricht sind dir besonders aufgefallen?</i> (nur berücksichtigt wenn SWSF02 mit mindestens "stimme eher zu" bewertet wurde)															
Gruppe	Disziplin/Atmosphäre	Unterricht ist fokussierter	Lehrperson-Schülerinnen und Schüler-Verhältnis verbessern	Besser erklärt	Unterrichtstempo/Wiederholungen	allgemein bemüht	mehr Experimente	konkrete Kritikpunkte aus dem Feedback verbessert	Abwechslungsreicher	Schülermitbeteiligung	Struktur	mehr Humor	Lehrperson hat sich über Ergebnisse lustig gemacht	sonstiges	gesamt
1	0,280	0,000	0,000	0,000	0,040	0,000	0,000	0,040	0,000	0,000	0,000	0,000	0,000	0,080	0,440
2	0,000	0,000	0,000	0,043	0,022	0,043	0,043	0,000	0,022	0,022	0,000	0,022	0,022	0,043	0,283
3	0,056	0,014	0,000	0,028	0,028	0,014	0,042	0,014	0,000	0,000	0,000	0,000	0,000	0,085	0,282
4	0,014	0,000	0,000	0,042	0,014	0,000	0,000	0,000	0,000	0,014	0,014	0,000	0,000	0,056	0,155
5	0,231	0,115	0,038	0,038	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,077	0,500

Es lässt sich also vorerst festhalten, dass in den beiden Gruppen, die zu Beginn die niedrigsten Werte in der Unterrichtsqualität zeigten, von den Schülerinnen und Schüler vor allem eine Verbesserung der Disziplin festgestellt wurde. Es wäre zu erwarten, dass sich dies auch in einer Verbesserung des *Klassenmanagements* im Prä-/Postvergleich der beiden Gruppen (Klassen) widerspiegelt. Dies ist jedoch nur in Gruppe 1 (Klasse 10) der Fall (siehe Abschnitt 5.3.1, Tab. 26).

Die beiden weiteren offenen Antwortformate (Items V2OFRA05 und V2OFRA06, allgemeine Beurteilung des Feedbackverfahrens) werden in den Tabellen 31 und 32 dargestellt. Auch hier

sind die relativen Häufigkeiten aufgeführt. Betrachtet man Tabelle 31 zeigt sich ein erwartungsgemäßer kontinuierlicher Anstieg der negativer Antwortkategorien mit sinkender Effektstärke im Prä-/Postvergleich. Es wird also der Trend des Prä-/Postvergleichs der Unterrichtsqualitäts-Entwicklung abgebildet.

Zwar ist dies bei den positiven Nennungen (Tab. 32) nicht ganz so eindeutig wie bei den negativen, trotzdem kann man von einem unerwarteten Ergebnis reden, da hier die Anzahl der positiven Nennungen mit sinkender Effektstärke ansteigen. Insbesondere die Gruppe 5 sticht hier heraus, die außerdem eine überdurchschnittliche Anzahl von Antwortkategorien aufweist, die sich auf die *positive Umsetzung der geäußerten Kritik* (Spalte 10) bezieht. Dieses überraschende Ergebnis relativiert sich jedoch bei erneuter Betrachtung von Tabelle 31. Hier wird deutlich, dass die Schülerinnen und Schüler der Gruppe 5 sich überdurchschnittlich oft beklagen, dass zu wenig am Unterricht geändert wurde. Die Wahrnehmung zwischen den Schülerinnen und Schüler bezüglich der Umsetzung der kritisierten Unterrichtsinhalte scheint in dieser Gruppe demnach zum Teil gegensätzlich zu sein.

Ein weiterer auffälliger Punkt in Gruppe 5 ist der Eindruck mancher Schülerinnen und Schüler, dass das Feedbackverfahren von manchen Mitschülern nicht ernst genommen wurde. Insgesamt zeigt sich auch, dass die Anzahl der negativen Nennungen pro Schüler in Gruppe 5 ca. zwei bis fünf mal so hoch ist wie in den anderen Gruppen, während bei den positiven Nennungen die Anzahl nur etwa um das 1.3 bis zweifache höher ist. Die erhöhte Anzahl negativen der Nennungen pro Kopf in Gruppe 5 kann also nicht allein damit erklärt werden, dass in dieser Gruppe generell häufiger das offene Antwortformat genutzt wurde. Es scheint eher umgekehrt zu sein, die insgesamt häufigere Nutzung des offenen Antwortformats in der Gruppe 5 trägt zum überraschenden Trend der positiven Antwortkategorien bei.

Tabelle 31: Allgemeine Beurteilung des Verfahrens, Item V2OFRA06

So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich schlecht, weil....														
Gruppe	nichts geändert	zu wenig geändert	Antwortmöglichkeiten zu ungenau	Mitschüler nehmen es nicht ernst	zu viel Zeit geht verloren	Die Ergebnisse fielen immer gleich aus	Fragen innerhalb des Bogens wiederholten sich	Die Lehrperson nahm es nicht ernst	zu wenig Zeit zum Ausfüllen	Fragebogen immer gleich	macht kein Spaß	Zusammenfassung der Individualurteile zum Klassenurteil	sonstiges	gesamt
1	0,000	0,000	0,000	0,000	0,040	0,000	0,000	0,000	0,000	0,000	0,040	0,000	0,080	0,160
2	0,000	0,022	0,000	0,043	0,043	0,000	0,000	0,000	0,000	0,022	0,022	0,000	0,087	0,239
3	0,014	0,028	0,028	0,014	0,085	0,000	0,014	0,000	0,014	0,028	0,014	0,028	0,056	0,324
4	0,042	0,042	0,014	0,028	0,113	0,014	0,014	0,014	0,042	0,000	0,028	0,000	0,042	0,394
5	0,038	0,192	0,000	0,192	0,115	0,000	0,000	0,000	0,038	0,000	0,000	0,000	0,231	0,808

Tabelle 32: Allgemeine Beurteilung des Verfahrens, Item V2OFRA05

So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich gut, weil....																
Gruppe	Meinung äußern allgemein	negative Kritik äußern können	positive Kritik äußern können	Kommunikation zwischen Schülerinnen und Schülern und Lehrperson verbessert	Anonymität hilft	Allgemeines Stimmungsbild der Klasse sehen	Lehrperson helfen den Unterricht zu verbessern	Schülerinnen und Schüler werden zur Reflektion des Unterrichts angeregt	Die Lehrperson die Kritik umsetzt	Unterrichtszeit wird weniger	Abwechslung vom Alltag	Fand es lustig	Gut für das Klassenklima	Es hat Spaß gemacht	sonstiges	gesamt
1	0,160	0,040	0,040	0,000	0,000	0,040	0,160	0,000	0,080	0,040	0,000	0,040	0,040	0,040	0,160	0,840
2	0,152	0,043	0,065	0,065	0,109	0,087	0,261	0,065	0,130	0,065	0,000	0,022	0,022	0,022	0,065	1,174
3	0,254	0,070	0,056	0,042	0,155	0,127	0,352	0,056	0,099	0,042	0,014	0,000	0,014	0,014	0,042	1,338
4	0,169	0,085	0,056	0,014	0,070	0,141	0,310	0,056	0,042	0,070	0,014	0,000	0,014	0,028	0,056	1,127
5	0,308	0,192	0,000	0,000	0,038	0,038	0,346	0,077	0,269	0,038	0,000	0,038	0,038	0,000	0,115	1,500

Generell bleibt festzuhalten, dass in der allgemeinen Beurteilung des Feedbackverfahrens das Muster der berechneten Gruppen teilweise wiederzufinden ist. So ist vor allem auffällig, dass die Anzahl der Nennungen innerhalb der negativen Antwortkategorien das ähnliche Muster aufweist, wie die Entwicklung der Unterrichtsqualität von der Prä- zur Posterhebung.

### 5.4.3 Intergruppenvergleich, Fragebogenergebnisse der Lehrpersonen

Der Intergruppenvergleich aus Sicht der Lehrperson (Fragebogen) basiert auf den Einschätzungen der Lehrpersonen hinsichtlich der *sozialen und kognitiven Unterrichts-Voraussetzungen* und der Einschätzung des *Beurteilungsvermögens der Schülerinnen und Schüler bezüglich Unterricht*. Diesbezügliche Daten, die in Tabelle 33 dargestellt sind, wurden zum Zeitpunkt der Präerhebung erfasst. Entwicklungen von der Prä- zur Posterhebung sind innerhalb dieser Skalen und Items kaum erkennbar und bringen keinen weiteren Erkenntnisgewinn.

Da die Lehrperson eine zentrale Rolle in der Unterrichtsentwicklung einnimmt, können die folgenden Daten nicht nur zur Untermauerung der gefunden Gruppen dienen sondern teilweise einen Beitrag zur Aufklärung der Ursachen für die unterschiedliche Entwicklung der Unterrichtsqualität leisten.

Interessanterweise spiegelt die Präerhebung der Einschätzung der *sozialen und kognitiven Voraussetzungen* und des *Beurteilungsvermögens der Schülerinnen und Schüler bezüglich Unterricht* nicht das Ergebnis der Präerhebung aus Schülersicht wider, sondern eher die Entwicklung der Unterrichtsqualität von Prä zu Post (Ausnahme: *Klassenmanagement*, siehe Tab. 33). Durchgängig die niedrigsten Werte verzeichnen die Gruppen 4 (eine Skala) und 5 (drei Skalen

und ein Item). Demzufolge ergibt sich, dass bei den Lehrpersonen, die in dieser Studie das *Beurteilungsvermögen der Schülerinnen und Schüler bezüglich Unterricht* als niedrig, die *Lernvoraussetzungen* als ungünstig, das *Schüler-Lehrpersonen-Verhältnis* als weniger gut und die Klasse als *leistungsmäßig heterogen* einschätzten, nur gering positive bis negative Entwicklungen der Unterrichtsqualität nach der Intervention aufweisen. Dieser Befund wird in die differenzierte Hypothesenbildung (Abschnitt 5.5.6 und Unterkapitel 6.4) zur Wirkung und zu den Wirkmechanismen von Schüler Feedback einfließen.

Tabelle 33: Einschätzung der kognitiven und sozialen Voraussetzungen, des Schüler-Lehrpersonen-Verhältnisses und des Beurteilungsvermögens der Schülerinnen und Schüler von Unterricht, Präerhebung

	Einschätzung der Lehrper- son: Schüle- rinnen und Schüler sind kompetente Beurteiler von Unterricht; je dunkler, desto kompetenter.	Einschätzung günstige Lern- voraussetzung Schülerinnen und Schüler; je dunkler, desto besser.	Schüler- Lehrpersonen- Verhältnis aus Lehrperson- Sicht; je dunk- ler, desto besser.	Gutes Klas- senmanage- ment aus Lehrperson- Sicht; je dunk- ler, desto besser.	Eingeschätzte Homogenität der Klasse aus Lehrperson- Sicht, je dunk- ler, desto homogener.
Gruppe 1	3.20	2.40	3.50	2.25	3.00
Gruppe 2	3.40	3.00	3.25	3.13	2.00
Gruppe 3	3.33	2.80	3.42	3.33	2.00
Gruppe 4	2.60	2.33	2.83	2.67	1.67
Gruppe 5	3.00	2.00	2.00	1.50	1.00

#### 5.4.4 Zusammenfassung der Intergruppenvergleiche

Die identifizierten Gruppen, die sich aus der Entwicklung der Unterrichtsqualität von der Prä- zur Posterhebung ergaben, wurden in der Analyse der weiteren Daten häufig wiedergefunden. Dies gilt in besonderem Maße für die Gruppe 1, die analog zu ihren unterdurchschnittlichen Werten zur Präerhebung und ihrem starken Anstieg zur Posterhebung unter anderem ähnliche Verläufe innerhalb der Kurzfeedbackintervention aufweist. Ebenso gilt dies, wenn auch nicht in dieser Eindeutigkeit, für die Gruppen 2 bis 4. Einen Sonderstatus nimmt die Gruppe 5 ein, die sich zwar klar von den anderen Gruppen abgrenzt, jedoch häufig nicht erwartungskonforme Werte in den anderen Daten zeigt. So entwickelten sich die Verläufe der Kurzfeedbacks, im Gegensatz zur betrachteten Unterrichtsqualität konstant bis positiv. Ebenso war das Antwortverhalten bei den offenen Frageformaten häufig widersprüchlich. Die Schülerinnen und Schüler dieser Gruppe äußerten sich überdurchschnittlich häufig positiv zum Feedbackverfahren und beschrieben am häufigsten beobachtete Veränderungen im Unterricht. Die Gruppen 1- 4 verhielten sich in letzterem Fall erwartungskonform.

Ein generell gutes Abbild des gefundenen Musters zeigen außerdem das offene Antwortformat bezüglich negativer Kritik am Feedbackverfahren auf Schülerseite (analog zu den Effektstär-



ken), und die Selbsteinschätzung zum Klassenmanagement (analog zur Prämessung der Unterrichtsqualität).

Die Ergebnisse des letzten Unterkapitels werden in Abschnitt 5.5.6 zur vorläufigen Beantwortung der Forschungsfrage 5 zusammengefasst.

### 5.5 Analyse der einzelnen Gruppen

Nun werden die im vorigen Unterkapitel identifizierten Gruppen unter zu Hilfenahme der Interviews (siehe auch Anhang 8.4) eingehender analysiert. Dabei wird insbesondere versucht, charakteristische Muster in der Unterrichtsentwicklung herauszuschälen beziehungsweise bereits identifizierte Muster mit den Informationen aus den Interviews zu konsolidieren, präzisieren, differenzieren beziehungsweise zu untermauern. Schließlich werden alle weiteren erhobenen Daten hinzugezogen, um eine abschließende Charakterisierung der einzelnen Gruppen zu erhalten, welche die Daten aus allen verfügbaren Instrumenten kombiniert und vernetzt.

Vor der qualitativen Analyse der Interviewergebnisse bezüglich der Unterrichtsentwicklung, wird die entsprechende relative Anzahl der Kategorien (Antworthäufigkeit pro Lehrperson) im Gruppenvergleich vorgestellt. Während der Kodierung der Interviews wurde zwischen Veränderung der Unterrichtsentwicklung hinsichtlich allgemeiner Wirkung des Schüler-Kurzfeedbacks und der Wirkung aufgrund konkreter Ergebnisse (*stimulated recall*) unterschieden. Da diese Unterscheidung nur schwer klar zu trennen ist und keine vertieften Erkenntnisse liefert, wurden die beiden Kategorien in Tabelle 34 zusammengefasst.

Zwei Auffälligkeiten zeigen sich: Die Lehrperson der Gruppe 1, die eine hohe Effektstärke in der Veränderung der Unterrichtsqualität im Prä-/Postvergleich aufweist, berichtet in den meisten Kategorien verhältnismäßig häufig von Veränderungen der Unterrichtshandlungen, beziehungsweise von Ausprägungen im Interpretations- und Motivationsprozess (siehe Abschnitt 5.5.1), die eine Handlungsänderung begünstigen (die beiden höchsten Werte sind grün markiert). Wohingegen Aussagen, die einer Handlungsänderung entgegenwirken (die beiden höchsten Werte rot markiert), eher von Lehrpersonen getroffen wurden, deren Gruppen geringere Effektstärken aufwiesen (Gruppe 5 und z. T. Gruppe 4).

Ein interessanter Aspekt nimmt die Ausprägung der Rezeption der Feedbackergebnisse ein. Mit sinkender Effektstärke steigen die Aussagen, das Feedback sachlich aufgenommen zu haben. Dagegen zeigt die Gruppe mit der höchsten Effektstärke die meisten Antworthäufigkeiten hinsichtlich der positiven Emotionen. Bei der Gruppe mit den niedrigsten Effektstärken verhält es sich diesbezüglich umgekehrt. Allerdings sei eingewendet, dass es sich bei den beiden genannten Gruppen um Einzelfälle handelt und somit immer nur eine Lehrperson in die Betrachtung eingeht. Hinsichtlich der negativen Emotionen ist kein Hinweis auf ein Muster zu erkennen. Es

scheint, als ob eine emotional neutrale Rezeption die weitere Motivbildung am wenigsten begünstigt.

Tabelle 34: Antworthäufigkeiten der Kategorien in den Interviews pro Lehrperson innerhalb der Gruppe

Unterrichtsentwicklung	Rezeption			Interpretation			Motivation		Handlung		Motivation für zukünftiges Unterrichtshandeln
	emotional positiv	emotional negativ	sachlich	Schüleraussage richtig aber nicht veränderbar	Schüleraussage richtig und veränderbar	Schüleraussage falsch	Änderungsabsicht vorhanden	keine Änderungsabsicht	Handlung verändert	Handlung unverändert	Beschreibung der Handlungsabsicht
Gruppe 1	4,00	0,00	2,00	2,00	3,00	0,00	7,00	0,00	6,00	0,00	1,00
Gruppe 2	1,50	0,50	4,00	3,00	3,00	0,50	3,00	2,00	2,50	3,00	1,50
Gruppe 3	0,67	0,67	3,00	2,33	2,00	0,33	3,00	2,00	2,33	0,33	1,33
Gruppe 4	2,33	2,67	7,33	4,00	2,00	0,67	5,00	0,33	4,00	0,67	1,67
Gruppe 5	0,00	0,00	8,00	5,00	3,00	1,00	2,00	2,00	2,00	3,00	0,00

### 5.5.1 Gruppe 1 (Klasse 10)

Wie soeben angedeutet, beschrieb die Lehrperson dieser Gruppe überdurchschnittlich häufig Handlungsänderungen aufgrund der Intervention, beziehungsweise nannte Aspekte, die eine Handlungsänderung begünstigen (die Lehrperson rezipierte die Feedbacks durchweg emotional positiv, attribuierte die Kritikpunkte vielfach als kontrollierbar und variabel; und es bildeten sich häufig Motivationen, das Unterrichtshandeln zu ändern). Die gebildeten Motivationen führten oft zu tatsächlichen Veränderungen im Unterrichtshandeln. Dazu gehörten häufigere Durchführungen von Versuchen, eine bessere und kleinschrittigere Strukturierung des Unterrichts, ein früheres Eingreifen bei Unterrichtsstörungen (Abschnitt 5.2.7.2, S. 80) und das Einbauen von Wiederholungsstunden. Der folgende, etwas längere Interviewausschnitt, verdeutlicht die Veränderungen und die unterschiedlichen Quellen für deren Motivationen:

*I [...] hat sich denn irgendwie eine Motivation ergeben, während der Feedbackphase etwas zu ändern, generell, am Unterricht jetzt so bei Ihnen?*

**B** [...] habe ich mir wirklich überlegt, was kann ich praktisch auch zeigen[...], jetzt steht in dem Feedbackbogen ja, beurteile das Praktische, Vorgehensweise, ehm da habe ich mir dann wirklich, mehr Gedanken im Vorfeld gemacht.

**I** Ja, ja, okay, Sie haben im Vorfeld sich mehr Gedanken gemacht und von den Ergebnissen? Eh, oder aufgrund der Ergebnisse, haben Sie sich da noch mal Gedanken gemacht, weil Sie jetzt, keine Ahnung, das erste Mal, das Ergebnis jetzt irgendwie nicht so war, wie sie jetzt, sich selber vorgestellt haben oder so, könnte ja sein. Also haben Sie sich aufgrund der Ergebnisse neu Gedanken gemacht zum Unterrichten?

**B** Nein, [...] die Schüler haben das schon genau so [...], im Feedbackbogen wiedergegeben haben, haben sie auch mit mir kommuniziert, das ist das Schöne an so einer lebhaften Klasse, dass sie halt ihr Feedback an den Lehrer geben. [...], ja ich bin etwas, ah wie soll ich sagen, doch sensibler an den Unterricht drangehen, ehm weil mir es doch im Hinterkopf hatte so, zum Schluss gibt es eine Befragung, wie wird man im Unterricht gesehen.

**I** Ja, ja, okay, mhm (<bejahend>) und wie haben Sie dann die Motivation umgesetzt? Also, man kann ja sagen, ich habe eine Motivation aber ehm, dass das wirklich in Handlung mündet, ist ja wieder was anderes, also was haben Sie generell, so richtig konkret gemacht, dann, verändert? Am Unterricht?

**B** Ehm, (.) Kritikpunkt war die Disziplin, die Lautstärke, also das habe ich dann, eh Umsetzten von bestimmten Schüler oder durch eh schnelleres, konsequenteres Handeln, ehm geändert, dann habe ich natürlich auch versucht, ehm Versuche reinzubringen, ein bisschen an der Methodenvielfalt eh zu ändern. Und auch eh, weil es gewünscht war durch die Schüler. Ehm, das ist interessant, wenn die im normalen Unterricht gefragt hätten, könnten wir das nochmals wiederholen. Und dann hätte ich gesagt, wir haben eine Einführungsstunde gemacht, wir haben eine Übungsstunde gemacht, wir haben eine Vertiefungsstunde gemacht, nee, also jetzt müsst ihr es irgendwie auf die Reihe bringen, aber als ich es dann wirklich so schwarz auf weiß hatte, dass es sich wirklich ehm, doch mehrere Schüler gewünscht haben, nochmal ehm eine Übungsstunde, oder nochmal wirklich eine Schritt für Schritt-Erklärung für das Thema und dann habe ich das gemacht, und dann sind natürlich wieder so so ein Teil der eh nicht aufpasst, und dann habe ich gesagt, so jetzt kucken wir mal, [...], hier, es haben sich Schüler gewünscht, das ist nicht einfach dahergeredet, sondern die haben es schriftlich wiedergegeben, wir wünschen uns nochmals eine Übungsstunde, oder eine Übungseinlage, oder so und dann haben selbst die die ehm, Störenfriede, Aktive, die, was weiss ich, ne, (<unverständlich>) die haben dann auch, ah ja, wenn das

*jetzt wirklich, das steht ja da, man hat es richtig gemerkt, das ist schwarz auf weiss, es ist gewünscht, und dann lief das, also (<unverständlich>).*

Zwei Aspekte sind bemerkenswert. Erstens, dass sowohl das Wissen über eine schriftliche Rückmeldung durch die Schülerinnen und Schüler (zu Beginn des Abschnitts) als auch die Rückmeldungen selbst motivierend für eine Handlungsänderung wirkten (letzter Abschnitt). Und zweitens, dass den Ergebnissen aus der schriftlichen Befragung auf Seiten der Lehrperson und auf Seiten der Schülerinnen und Schüler eine höhere Bedeutung zugemessen wurde als ein mündliches Feedback von den Schülerinnen und Schülern. Somit konnten Handlungsänderungen, die im Unterricht durchgeführt wurden, besser legitimiert werden.

In einem weiteren Beispiel zeigt sich jedoch auch, dass nicht alle erfolgreichen Handlungsänderungen (hier verbesserte Struktur innerhalb des Unterrichts) aufgrund des erforderlichen Arbeitsaufwandes nachhaltig durchgeführt werden können:

***B** Ja, ja, ja, bei dem weiß ich es, das war, das war wirklich eh spaßig, da hatte ich dann gesagt, genaue Uhrzeit aufgeschrieben, ehm, Einführungsphase von eh 14Uhr bis 14Uhr20, Übungsphase bis zum ersten Klingen 14Uhr45, dann nochmal fünf Minuten, eh, Kurzfrage und dann schreiben wir den Test. Und die Schüler fanden das Klasse, die haben dann gesagt, so wollen wir immer unseren Unterricht.*

***I** Ja*

***B** Hehehehe (<lacht>)*

***I** Gut, ja machen Sie es jetzt auch immer, so?*

***B** Ehm, nee weil das war wirklich eh, viel Vorbereitungsphase für die ganze ehm Übungsblätter, und und so weiter. Also da habe ich mich wirklich, das war jetzt weniger wegen der Umfrage, sondern da wollte ich, ehm für die Stunde, so und so ein bisschen mehr Struktur haben. [...]*

Die Lehrperson passte den Unterricht also auf mannigfaltige Weise an, sieht aber in der nachhaltigen und langfristigen Änderung durch das Schüler-Feedback eher wenig Chancen und gibt auch nicht an, Motivationen für zukünftiges Unterrichtshandeln gebildet zu haben.

Trotzdem würde sie ein Schüler-Feedbackverfahren, wenn es in elektronischer Form verfügbar wäre, ein bis zweimal im Jahr einsetzen. Außerdem sah sie das Schüler-Lehrpersonen-Verhältnis und die Disziplin in der Klasse größtenteils verbessert. Diese Aussagen decken sich mit denen der Schülerinnen und Schüler aus der Prä- und Posterhebung.

### **5.5.1.1 Charakterisierung Gruppe 1**

Die Gruppe 1 (Klasse 10) lässt sich also wie folgt charakterisieren: Die Werte der Unterrichtsqualität aus Schüler-Sicht waren zu Prä die zweittiefsten im Vergleich zu den anderen Gruppen. Allerdings entwickelte sich diese Gruppe im Verlauf der Intervention mit der stärksten positiven Tendenz. Dies wird sowohl bei den Verläufen der Skalen innerhalb der Kurzfeedbacks sichtbar (vier von fünf Skalen entwickeln sich signifikant positiv) als auch in der Prä-/Posterhebung der Unterrichtsqualität (Durchschnittliche Effektstärke über alle Skalen  $d = 0.57$ ).

Die Schülerinnen und Schüler sehen in dem Verfahren kaum Nachteile und bemerken vor allem, dass sich die Disziplin/Unterrichtsatmosphäre verbessert hat. Die Lehrperson dieser Gruppe beurteilte im Vergleich zu den anderen Gruppen das Schüler-Lehrpersonen-Verhältnis am besten und schätzte die Klasse leistungsmäßig am homogensten ein. Obwohl die Lehrperson die Beurteilungsfähigkeit der Schülerinnen und Schüler im Vergleich zu den anderen Gruppen durchschnittlich einstuft (die Werte lagen jedoch bei jeder Lehrperson über dem mittleren Wert von 2,5), nahm sie die rückgemeldeten Kritikpunkte sehr ernst und veränderte auf vielfältige Weise den Unterricht. Sie baute mehr Experimente ein, verbesserte die Struktur, schob Übungsstunden ein und intervenierte früher bei Unterrichtsstörungen. Die Lehrperson hob dabei hervor, dass es gerade die schriftliche Form des Feedbacks war, die zur Motivationsänderung und zur Legitimierung der Handlungsänderungen, auch vor den Schülerinnen und Schülern, beitrug.

Dieses insgesamt positive Bild wird jedoch auf zweifache Weise eingeschränkt. Zum einen lag die Motivation zur Veränderung des Unterrichtshandelns nicht nur an den Rückmeldungen der Schülerinnen und Schüler selbst, sondern auch in der Erwartung, dass die Stunde bewertet wird. Zum anderen betonte die Lehrperson, dass sie den Arbeitsaufwand, den sie während der Feedbackphase leistete, nicht weiter leisten könne und somit die Nachhaltigkeit des Verfahrens anzweifelte. Trotzdem beurteilte die Lehrperson das Verfahren insgesamt positiv und sah sich motiviert, in Zukunft öfters ein Schülerfeedback durchzuführen.

### **5.5.2 Gruppe 2 (Klassen 7 und 8)**

Analysiert man die Interviewaussagen bezüglich der Reaktionen auf das Schüler-Kurzfeedback (Tab. 34), fällt auf, dass die Antworthäufigkeiten nur bedingt erwartungskonforme Werte annehmen. Auf Ebene der handlungsbegünstigenden Sub-Kategorien sind sie eher tief, und auf Ebene der handlungsverhindernden Sub-Kategorien weist die Gruppe gar zweimal die höchsten Werte auf (Siehe Tab. 34, rote und grüne Markierungen).

Erwartungskonform sind die Ergebnisse hingegen innerhalb der Kategorie *Interpretation*. Die rückgemeldeten Kritikpunkte werden meist als veränderbar und kontrollierbar attribuiert und entsprechen somit den Erwartungen. Die Art und Weise, wie die beiden Lehrpersonen auf die

Feedbacks reagierten, beziehungsweise welche Handlungen sie veränderten unterscheidet sich jedoch größtenteils. Während die Lehrperson der Klasse 8 vorrangig auf die offenen Antwortformate reagierte und inhaltliche Fragen wiederholt aufgriff, versuchte die Lehrperson der Klasse 7 die Experimentierphasen etwas stärker anzuleiten und veränderte somit eher methodische Aspekte. Gemeinsam war den beiden Lehrpersonen jedoch, dass auch sie, wie die Lehrperson der Gruppe 1, den Unterricht intensiver vorbereiteten, weil sie wussten, dass ein Feedback ansteht:

Lehrperson Klasse 8:

*I [...] genau also Motivation, jetzt zum Beispiel am Unterricht was zu ändern, hatten Sie da ehm, gab es das in dieser Feedback-Phase so ganz generell?*

*B [...] bevor ich die Stunde gemacht habe sozusagen, war das wie so ein, ein kleines Teufelchen, was einen selber beobachtet, vielleicht, kann man sagen, was hast du jetzt mit denen genau vor, ja, also sozusagen, nach der Planungsphase so mal noch eine extra Planungsphase, eingeschoben für mich, wo ich dann mal überlege, kann ich das jetzt so machen, wenn das dann nachher abgeprüft wird, ja, oder „Wenn das dann nachher abgeprüft wird“ ist falsch, wenn das dann nachher von den Schülern zurückgemeldet wird und dann auch in andere Kreise geht und so. Also von dem her würde ich sagen, habe ich schon über jede Stunde nochmal, aber erst am Anschluss eigentlich, wenn es mir dann wieder eingefallen ist, nochmal [...] darüber nachgedacht. Aber das ist jetzt nicht so, also ich unterrichte jetzt nicht zum ersten mal ne neunte Klasse, dass ich gesagt habe, ich schmeiße das über den Haufen und mache jetzt was Neues, sondern das ist schon so geblieben wie es ist. Wie schon gesagt, das war eher auf so einer, vielleicht eher auf einer motivationalen Ebene, dass das jetzt sowohl für die Schüler als auch mich glaube ich eine Bedeutung hatte.*

*I mhm (<bejahend>)mhm (<bejahend>) also aber, aufgrund der Rückmeldung von den Schülern haben Sie jetzt keine große Änderung?*

*B [(<Überlappung Anfang>) nein, habe ich jetzt keine Änderung gemacht](<Überlappung Ende>)*

Lehrperson Klasse 7:

*I Inwiefern haben Sie denn den Unterricht anders entwickelt, weil Sie wussten dass heute ein Feedback kommt?*

*B (...) Ich, ich hab mehr auf [...] Schüleraktivitäten halt eh geachtet, dass die halt eben mehr zu tun haben, eh das ist vielleicht auch ein bisschen, ehm kleinschrittiger leichter nachzuvollziehen ist, diese Sachen, und im Prinzip kleine abge-*

*schlossene Geschichten. Man ist immer im Physikunterricht verleitet, [...] man macht was, eh es läutet, und in der nächsten Stunde geht es dann irgendwie so weiter und da war es jetzt halt wirklich wichtig, dass jetzt halt wieder eine Stunde, eine Stunde ist eh, wo man was macht und wo man am Ende, man sagen kann, okay, ich habe heute was verstanden, ich habe heute was gelernt und so. Eh also dieser Rahmen der Stunde ist dann für mich wieder wichtiger geworden.*

*I mhm (<bejahend>)*

**B** *Sonst sollte man das ja eher versuchen, groß, dass man das eher als Gesamtprojekt das ganze Jahr sieht und dass eben hat natürlich auch den Nachteil, dass man halt dann auch ein bisschen fokussierter vorgehen muss und manchmal ne Frage halt dann abblocken muss und sagen, he wir haben jetzt noch so viel Zeit jetzt, um den Bogen auszufüllen, das sind dann halt die Nachteile von der ganzen Sache wieder aber ansonsten war es halt mehr so eine Geschichte.*

Beide Interviewausschnitte deuten darauf hin, dass die beiden Lehrpersonen aufgrund der bevorstehenden Feedbacks, sich vor allem strukturell intensiver Gedanken gemacht haben. Es scheint, als brächen sie partiell aus der Routine aus, gingen einen Schritt zurück und überlegen nochmals, was sie in der Stunde vorhaben und mit welchen Schritten sie das umsetzen wollen. Die Reflektion der geplanten Schulstunden bekommt also eine größere Bedeutung und orientiert sich stark an den Fragestellungen des Kurzfeedbacks. Einerseits weil die Lehrpersonen wissen, dass sie bewertet werden und die Ergebnisse vor sich selbst verantworten müssen aber auch weil die Daten in fremde Hände gelangen.

### **5.5.2.1 Charakterisierung Gruppe 2**

Die Gruppe 2 charakterisiert sich dadurch, dass sie bezüglich der Unterrichtsqualität aus Schülersicht von einem vergleichsweise mittleren Niveau gestartet ist und daraufhin die zweithöchsten Effekte erzielte. Eine Tendenz, die sich zum Teil in den Bewertungen der Schülerinnen und Schüler bezüglich des Feedbackverfahrens als auch in den Schülerkurzfeedbacks widerspiegelt. Ergebnisse aus den Schülerkurzfeedbacks findet man beispielsweise auch konkret in den Interview-Aussagen wieder. So bemerkte die Lehrperson der Klasse 8, dass sie den Fokus nicht auf das wissenschaftliche Arbeiten gelegt hat, obwohl die Werte im Vergleich zu den anderen Skalen eher tiefer lagen; die Steigung dieser Skala über die vier Messzeitpunkte weist im Vergleich zu den anderen Gruppen den konstantesten Wert auf.

Außerdem beurteilen die Lehrpersonen dieser Gruppe die Lernvoraussetzungen ihrer Schülerinnen und Schüler am besten und sie schätzen die Kompetenz der Schülerinnen und Schüler, Unterricht adäquat zu beurteilen, am höchsten ein. Ein Standpunkt, der sich größtenteils in der Interpretation der Schüleraussagen wiederfindet, die nur in seltensten Fällen als falsch beurteilt

wird (siehe Tab. 34). Die Motivationen zu Veränderungen in den Unterrichtshandlungen finden, ebenso wie in Gruppe 1, nicht nur ihren Ursprung in den Schüleraussagen, sondern auch in der Erwartung, dass ein Schüler-Feedback bevorsteht.

Die berichteten Änderungen im Unterrichtshandeln sind jedoch in ihren Ausprägungen weniger mannigfaltig als in Gruppe 1. Interessant und kaum erwartungskonform ist, dass die Lehrperson der Klasse 7 die Feedbacks nur vorgestellt, aber nie mit der Klasse über die Ergebnisse diskutiert hat, weil sonst die Anonymität verloren gegangen wäre. Auch die Lehrperson der Klasse 8 bemängelte, dass durch die Diskussion die Anonymität verloren gehe und sie empfand die Diskussionen als sehr träge; sie stellte fest, dass die Schülerinnen und Schüler nicht das Bedürfnis hatten nochmals über die Ergebnisse zu reden. Man könnte aber auch die Hypothese aufstellen, dass die Diskussionen deshalb träge verliefen, weil die Schülerinnen und Schüler ihre Anonymität nicht aufgeben wollten. Beide Lehrpersonen würden das Verfahren in Zukunft – jedoch weniger geballt – einsetzen, wenn es auf freiwilliger Basis geschähe und die Daten vor Missbrauch geschützt wären.

### 5.5.3 Gruppe 3 (Klassen 2, 4 und 6)

Die Anzahl der Nennungen bezüglich der Unterrichtsentwicklung aufgrund des Feedbacks sind in dieser Gruppe vergleichsweise durchschnittlich bis unterdurchschnittlich (Tab. 34). Dies gilt sowohl für handlungsfördernde als auch für handlungshemmende Aspekte. Da die Werte der Präerhebung dieser Gruppe die höchsten, die Effektstärken im Vergleich zu den anderen Klassen eher durchschnittlich waren und gleichzeitig wenig negative Kritik von den Schülerinnen und Schülern kam (siehe Abschnitt 5.4.1.7) entsprechen diese Werte den Erwartungen und passen somit ins Muster.

Kam es zu Handlungsänderungen, so fielen diese in den einzelnen Schulklassen unterschiedlich aus. Während die Lehrperson der Klasse 2 versuchte, mehr selbständig arbeiten zu lassen, baute die Lehrperson der Klasse 4 mehr Experimente in den Unterricht ein. Die Lehrperson der Klasse 6 versuchte den Fokus auf das wissenschaftliche Arbeiten zu richten und den Schüleraussagen in Klassengesprächen mehr Gewicht zu geben. Interessant war in dieser Gruppe, dass nur eine Lehrperson angab, den Unterricht aufgrund des Wissens, dass ein Feedback folgen werde, veränderte. Die Lehrperson der Klasse 6 berichtete, dass die Schulstunden, die von den Schülerinnen und Schülern rückgemeldet wurden, möglichst ein zentrales Lernziel beinhalten sollten:

*I Mhm (<bejahend>)mhm (<bejahend>) okay, ehm, inwiefern haben Sie denn den Unterricht anders entwickelt, weil Sie wussten, dass es heute ein Feedback gibt?*



**B** (..) Hmm, also ich habe, ja eigentlich gar nicht so viel verändert, was mir wichtig war, wobei, das ist mir in Physik immer wichtig, (<lacht>) unabhängig jetzt davon ob ein Feedback oder nicht, aber was mir natürlich trotzdem auch in der Stunde wichtig war, ist, dass die Schüler, was mit nach Hause nehmen, also irgend, irgendwie einen physikalischen Gedankenblitz hatten, und jetzt wissen, aha, deswegen ist das so und so, so findet man das im Alltag, so findet man auch die technische Umsetzung. Das ist mir einfach wichtig. Also, die sollen nicht das Gefühl haben, ehm ich lerne nur, für die Schule sondern ich habe vielleicht selber jetzt auch durch dieses Hintergrundwissen, ist es selber für mich ein Gewinn, diese Erkenntnis.

**I** Ja, ja, ja, also haben Sie den Unterricht ein bisschen anders entwickelt, wenn jetzt kein Feedback gewesen wäre, hätten Sie da, hätten Sie was.

**B** Da wäre es mir für, ja da wäre es mir vielleicht nicht so wichtig gewesen, dass es genau in der Stunde auch zu dem Ziel kommt, und da war es mir dann wichtig, dass es in der Stunde, auch zu dem Ziel kommt. Ansonsten hätte ich das vielleicht auch in einer anderen Stunde verlegt oder, mhm (<bejahend>).

Auch dies lässt sich dahingehend interpretieren, dass der Stunde eine stärkere Strukturierung gegeben wurde und dass die Lehrperson von der Alltagsroutine abwich, um sich nochmal genau zu überlegen, was in der Stunde dran kommen soll und wie man zum jeweiligen Unterrichts-/Lernziel gelangt.

### **5.5.3.1 Charakterisierung Gruppe 3**

Die Gruppe 3 zeichnet sich also dadurch aus, dass sie zu Prä die höchsten Werte in der Beurteilung der Unterrichtsqualität aufzeigte und sich mit kleinen Effekten vergleichsweise durchschnittlich entwickelte. Auch hier finden sich die Muster zum großen Teil in den Schülerfeedbacks wieder, die erstens auf hohem Niveau verlaufen und zweitens wenig negativ-kritische Rückmeldungen beinhalten und somit selten Ansatzpunkte zur Weiterentwicklung des Unterrichts enthalten. Dies spiegelt sich in den wenigen Handlungsänderungen während der Feedbackphase wieder.

Ein weiterer auffälliger Aspekt in dieser Gruppe sind die Einschätzungen der Lehrpersonen bezüglich der sozial-kognitiven Voraussetzungen und der Beurteilungsfähigkeit von Unterricht der Schülerinnen und Schüler. Diese nehmen in allen Skalen den höchsten oder zweithöchsten Wert an (siehe Kap 5.4.3, Tab. 33). Ob allerdings die positivere Einschätzung der Lehrperson oder die tatsächlich besseren Voraussetzungen für die positiven Ergebnisse der Unterrichtsqualität durch die Schülerinnen und Schüler zu erklären ist, lässt sich daraus nicht ableiten. Auf den Punkt gebracht ließe sich die Gruppe 3 mit der Aussage „Die hohe Einschätzung beruht auf

Gegenseitigkeit und führt zu stabilem Unterrichtsgeschehen auf hohem Niveau“ charakterisieren.

#### 5.5.4 Gruppe 4 (Klassen 1, 5 und 9)

Die Lehrpersonen der Gruppe 4 nennen im Durchschnitt die meisten negativen Emotionen während der Rezeption und die zweitmeisten external-stabilen/unkontrollierbaren Attributionen während der Interpretation der Schüler-Feedbacks („Schüleraussage richtig aber nicht veränderbar“). Ebenso geben sie am zweithäufigsten an, dass die Schülerurteile falsch seien (allerdings auf tiefem Niveau). Trotzdem wurde in dieser Gruppe häufig von Motivationen und tatsächlichen Handlungsänderungen berichtet. Lehrperson 1 und Lehrperson 5 haben das Unterrichtstempo angepasst, versucht die Unterrichtsatmosphäre zu verbessern, indem stärker auf Disziplin geachtet wurde und sie wollten sich mehr Zeit für Experimente nehmen. Lehrperson 9 änderte nichts am allgemeinen Unterrichtskonzept, hat aber auf konkrete Schüleraussagen reagiert, blieb jedoch insgesamt sehr vage und attribuierte sehr stark external:

*I Und der, haben Sie denn immer gewusst, beziehungsweise haben Sie Dinge gesehen, wo Sie denken, da könnte man was verändern?*

*B Mmm, also, mmm, wie gesagt, man kann natürlich in jeder Stunde noch mehr Experimente machen, man kann natürlich dann noch mehr denn noch eh Schülerzentrierter die Sachen dann nachher angehen, es liegt natürlich aber auch an der Ausstattung, es liegt natürlich auch an den Ressourcen, die vorhanden sind und von der Schule vorgegeben sind. Man muss eigene Produkte entwickeln und das Problem ist dann halt einfach mit der Sicherheit, so ein Sicherheits-Gefahrenrisiko was halt eben dort dann nachher schon wieder einschränkend ehm ist.*

*I Ja, ja mhm (<bejahend>) Und hat sich irgendwie eine Motivation entwickelt, irgendwas zu ändern, also, Sie haben jetzt grad gesagt, man könnte natürlich immer. Aber hat sich dann tatsächlich dann eine Motivation ergeben? Dass Sie jetzt was ändern wollten? Oder vielleicht noch wollen?*

*B Also natürlich ehm, besteht da immer eine Bestrebung oder so etwas zu verändern, mein, ich bin jetzt auch das erste Jahr hier an der Schule (<lacht>). Dann muss ich natürlich erstmal mit den ganzen Räumlichkeiten und auch mit den Klassen dann nachher arrangieren.*

*I Ja*

*B Und eh das natürlich auch schon sicherlich ein Thema, wo man dann sagt, ja man könnte vielleicht das eine oder andere noch besorgen, oder ja. Man müsste*

*halt auch die Physikstunden so gestalten, dass sie halt Doppelstunden wären. Wo man wirklich dann auch Sachen ausbauen kann, weil in einer einstündigen und ich habe immer Physik in der sechsten Stunde, eh ist das natürlich schon etwas anderes als wenn ich das in der ersten Stunde hätte also von dem her.*

*I Mhm (<bejahend>) und ehm gibt es da irgendwas konkretes, was Sie jetzt sagen können, was Sie jetzt ändern würden aufgrund dieser Ergebnisse?*

*B Mmh, schwierig, also man müsste es im Einzelnen oder so nochmals anschauen, ehm natürlich eh ist natürlich eh die Frage natürlich immer, es wurde ja, es wurde das Lernklima angesprochen, he, was versteht man jetzt eben quasi unter ehm super Lernklima, was versteht man unter mittlerem Lernklima, muss es halt dann auch ein bisschen abwägen können mit irgendwelchen Vergleichswerten und irgendwie dann nachher mit anderen Sachen vergleichen können, um da eine stichhaltige Aussage zu ermöglichen, ja.*

*I Ja, also gab es denn vielleicht auch tatsächlich eine Handlung? Oder haben sie etwas an Ihrem Unt., an Ihrer Unterrichtsentwicklung, oder am Unterricht selber dann während des Unterrichtens, während der Feedbackphase verändert? Oder ja?*

*B Also, nochmal, verändert, in dem Punkt, eigentlich, dass wir halt mehr über die einzelnen Sachen nochmals gesprochen hatten. Dass wir einfach nochmals, eh diskutiert hat oder so, was war denn da ganz gut oder was hat euch da denn nachher nicht so gut gefallen?*

*I Ja*

*B So in diese Richtung.*

*I Ja*

*B Das schon ja, aber jetzt ehm inhaltlich, ehm eigentlich weniger, ja.*

*I Inhaltlich und pädagogisch oder didaktisch auch nicht?*

*B Ehm (..) eigentlich eh, ne also nicht, jetzt, wie ich jetzt eigentlich schon bisher da entsprechend ehm das gemacht hatte.*

Erst als die Lehrperson im Verlauf des Interviews direkt auf konkrete Ergebnisse aus dem Schüler-Feedback angesprochen wurde (*stimulated recall*), beschrieb sie Motivationen und Handlungsveränderungen. Wie z. B., dass sie aufgrund der Rückmeldung „es sei zu laut“, die Sitzordnung veränderte; dass sie die Aspekte des wissenschaftlichen Arbeiten nochmals wiederholte und versuchte die Schülerinnen und Schüler mehr selbständig arbeiten zu lassen. Zudem hat die Lehrperson teilweise vorgängig den Unterrichtsverlauf angepasst, weil sie wusste, dass ein Schüler-Feedback bevorsteht.:

Lehrperson Klasse 9:

*I Mhm (<bejahend>) mhm (<bejahend>) Okay, eh hat, inwiefern haben Sie denn den Unterricht anders entwickelt, weil sie wussten, dass heute jetzt ein Feedback stattfindet?*

*B (.) Naja gut, mein eben so viel Zeit auf Grund meiner vielen Fächer hatte ich gar nicht um entsprechend groß zu intervenieren, habe eigentlich auch nicht gesehen, dass jetzt da der Bedarf, eh aufgrund der Ergebnisse, besonders groß sei in jetzt bestimmten Bereichen. [...] man hat natürlich dann auch schon ein bisschen geschaut welche Versuche eben, man dann nachher auswählt, für den Teil, weil ja vieles ja dann nachher auch am anderen Tag stattgefunden hat, da hat man halt auch noch ein bisschen geschoben, dass man gesagt hat, okay, für diese Zeiträume nimmt man jetzt ein besonders gutes Verfahren oder dergleichen, weil es halt einfach dann nachher nochmals eindrücklicher wird, [...]. Aber grundsätzlich bin ich eigentlich der normalen Ausarbeitung, eigentlich normal treu geblieben und habe jetzt nicht darüber hinaus jetzt einfach gesagt, ja jetzt heute machen wir mal jetzt irgendwie ein anderes Thema, obwohl ich dann nachher die Kisten rausholen kann und dann nachher, da punkten kann sondern die normale Fahrspur halt vom Unterricht habe ich belassen, ganz bewusst oder so um dann nachher zu sehen, naja wie reagieren sie jetzt auf das ganz normale um jetzt nicht irgendwie einen besonderen Effekt irgend um das Ergebnis zu verfälschen oder nochmal irgendwas anderes in den Vordergrund zu schieben, was es ja sicherlich auch gab, (<lacht>), deshalb.*

Man erkennt zwei Qualitätsunterschiede in den Veränderungsbemühungen der Unterrichtsentwicklung zu den Lehrpersonen der Gruppen 1-3. Erstens: Während letztere sich nochmals Gedanken um eine sinnvolle *Struktur* der Unterrichtsstunde machten, berichtet die Lehrperson der Klasse 9 davon, eindrücklichere Versuche von anderen Stunden in die jeweilige Stunde geschoben zu haben, um den Schülerinnen und Schülern den Inhalt des Unterrichts in der rückgemeldeten Stunde stärker zu verdeutlichen. Zweitens: Die Bemühungen, die Unterrichtsentwicklung anzupassen waren relativ gering, da die Lehrperson mit den Ergebnissen zufrieden war und keinen großen Handlungsbedarf sah.

#### **5.5.4.1 Charakterisierung Gruppe 4**

Zusammenfassend bleibt festzuhalten, dass Gruppe 4 bezüglich der Unterrichtsqualitätswerte die heterogenste aller Gruppen ist. Dennoch kann man von vergleichsweise mittleren Werten zu Prä und vergleichsweise unterdurchschnittlichen Effekten (bzw. nicht vorhandenen Effekten) sprechen. Dieses Muster findet sich auch in den Schülerfeedbackverläufen wieder, welche sich

auf vergleichsweise durchschnittlichem Niveau bewegen und in keiner Skala eine signifikante Steigung aufweisen.

Es fällt auf, dass die Lehrpersonen dieser Gruppe während der Interpretation der Schülerrückmeldungen eher ungünstig für eine Motivationsbildung attribuierten oder die Schülerurteile als falsch ansahen. Trotzdem wurden häufig Veränderungen im Unterrichtshandeln genannt. Diese unterschieden sich aber, passend zur Heterogenität der Gruppe, in ihrer Ausprägung. Während die Lehrperson der Klasse 9 ausschließlich auf konkrete Kritikpunkte reagierte und sonst kaum Handlungsbedarf sah, versuchten die Lehrpersonen der Klassen 1 und 5 zusätzlich pädagogische und didaktische Herangehensweisen zu verändern.

Des weiteren fällt auf, dass die Lehrpersonen dieser Gruppe im Vergleich zu den Gruppen 1-3 einem folgenden Feedback im Vorhinein eine geringere Bedeutung zumaßen und die Unterrichtsentwicklung vorgängig weniger veränderten. Außerdem schätzen die Lehrpersonen dieser Gruppe das *Beurteilungsvermögen der Schülerinnen und Schüler bezüglich Unterricht* am geringsten ein und die Skalen der *sozial-/kognitiven Voraussetzungen* der Schülerinnen und Schüler weisen die zweittiefsten Werte auf. Aber auch hier würden zwei von drei Lehrpersonen das Verfahren in Zukunft einsetzen, die Lehrperson der Klasse 5 würde dies im Falle eines vollautomatisierten Verfahrens sogar über mehrere Monate jede Stunde tun.

### 5.5.5 Gruppe 5 (Klasse 3)

Die Lehrperson der Gruppe 5 rezipierte die Rückmeldungen der Schülerinnen und Schüler ausschließlich sachlich, auch wenn es sich um positive Entwicklungen innerhalb der Intervention handelte. Dabei gab sie den Ergebnissen keine hohe Bedeutung, weil sie davon ausging, auch ohne systematische Rückmeldung gut erkennen zu können, ob der Unterricht gut oder schlecht lief (siehe Anhang 8.4). Die Lehrperson berichtete am häufigsten, dass die rückgemeldeten Kritikpunkte unveränderbar und die Schüleraussagen falsch sind (letzterer Punkt ist jedoch aufgrund der Tatsache, dass die Gruppe nur aus einer einzelnen Klasse besteht und die Anzahl der Nennungen insgesamt sehr gering ist, als bedingt aussagekräftig einzustufen). Außerdem erwähnte die Lehrperson am seltensten Motivationen zu Handlungsänderungen sowie deren Durchführungen. Falls doch, äußerten sich diese vor allem in Wiederholungen von nicht verstandenen Unterrichtsthemen. Die genannten Punkte kann man in Bezug auf die Entwicklung der Unterrichtsqualität als erwartungskonform werten, jedoch nicht auf die Verläufe der Kurzfeedbacks (siehe Abschnitt 5.4.1), deren Anstieg zu MZ 3 die Lehrperson folgendermaßen erklärt:

*B [...] Eh und ich habe da eine, eine Stunde in, in Erinnerung, die fällt dann eventuell aus dem Rahmen, das weiß ich jetzt nicht mehr, das war die dritte Erhe-*

*bung, und die ehem, war die erste wo die, wo die Zahlen angestiegen sind. Wo also im Nachhinein dann, dann "der Unterricht war nicht zu schwer", "nicht zu leicht", glaub ich beides mit drei oder besser herauskam und ehm, da habe ich mir natürlich eine ja wie man es halt macht, eine Erklärung zusammengereimt und mir überlegt woran kann das denn liegen, dass das jetzt auch immer besser war als vorher? Dann viel mir auf, ja, das war dann die Stunde wo ich, eh wo die Schüler experimentiert haben, erstens, und zweitens wo klar war, ich sammle das nachher ein und benote das und ehm, da eh sehe ich dann manchmal auch daran, ja das sind manchmal auch ganz andere Gründe, warum jetzt meine Schüler mitarbeiten und warum nicht und da lief das tatsächlich ganz gut. Die haben sich Mühe gegeben, die waren mehrheitlich ganz gut dabei und haben ehm, versucht, die Aufgaben zu lösen nach Kräften und ehm, ja die Ergebnisse waren dann so leidlich bei bei der Physikaufgabe und im Fragebogen dann deutlich besser als vorher. Ehm, und ich schiebe das dann auf diese Situation, dass sie wussten, es geht jetzt um etwas. Plötzlich, jetzt gibt es da eine Note darauf.*

Die Lehrperson führte die besseren Ergebnisse also einerseits darauf zurück, dass die Schülerinnen und Schüler mehr selbständig arbeiten konnten und dass die Arbeit benotet wurde. Wo bei sie letztgenannten Punkt stärker betont. Zu MZ 4 blieben die Werte größtenteils auf dem höheren Niveau, die Lehrperson findet hierfür aber keine stichhaltige Erklärung:

*I Mhm (<bejahend>) alle anderen Skalen blieben im Prinzip auf demselben Niveau, [...] aber ansonsten, ehm eh war die, waren die Ergebnisse von den Schülern, auf dem gleichen Niveau mehr oder weniger, wie zu Messzeitpunkt drei.*

**B** *Beim Vierten?*

*I Genau, bei Messzeitpunkt vier war es außer bei kognitiver Passung nicht zu schwer, eh mindestens gleich hoch.*

**B** *Mhm (<bejahend>) (..) ja, vielleicht, eh (..), vielleicht ist tatsächlich die Aufmerksamkeit ein bisschen gestiegen, insgesamt, und dann steigt auch alles andere, das ist, da gibt es glaube ich schon einen Zusammenhang. Diese Dinge sind ehm, der atmosphärische Teil ist ist eh natürlich angenehmer zu arbeiten wenn man was versteht und wenn man zuhört und wenn ich das nochmal erkläre, wenn da gefragt wird und ehm wenn wir uns da nicht gegenseitig anzicken und das ist vielleicht wenigstens bei einigen ein bisschen besser geworden. Ich glaube die Klasse zerfällt, es gibt so einen Gradienten zwischen denen, die versuchen, ganz viel mitzukriegen, viel beizutragen. Wir haben einige Schülerinnen*

*vor allem, also, mir fallen jetzt zwei ein und von den Jungen naja eigentlich nur einer, die wirklich regelmäßig gut mitarbeiten, [...].*

Die Aussage des letzten Satzes, passt gut zu einigen Äußerungen von Schülerinnen und Schülern, die den Eindruck hatten, dass ihre Mitschüler das Schüler-Feedback nicht ernst nähmen (siehe Abschnitt 5.4.2, Tab. 31). Auch in den erhöhten Varianzen dieser Gruppe innerhalb der Kurzfeedbacks und der beurteilten Unterrichtsqualität findet dies Bestätigung (Tab. 35 und 36). Dabei fällt auf, dass auch Gruppe 1 eher höhere Varianzen aufweist, obwohl beide Gruppen, im Gegensatz zu den anderen, nur aus einer Klasse bestehen und somit die Wahrscheinlichkeit einer höheren Varianz eher niedriger sein sollte, da zur Varianz beitragende Faktoren wie beispielsweise verschiedene Lehrperson und unterschiedlicher Unterrichtsstoff, wegfallen. Somit sind den, bereits erhöhten Werten, eine noch stärkere Bedeutung zuzumessen.

Tabelle 35: Varianzen der Schüler-Kurzfeedbacks in den Gruppen 1-5

<b>Durchschnittliche Varianz über alle Skalen der Schüler-Kurzfeedbacks</b>					
	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5
<b>MZ 1</b>	0,46	0,36	0,41	0,43	0,44
<b>MZ 2</b>	0,57	0,41	0,36	0,50	0,67
<b>MZ 3</b>	0,35	0,38	0,48	0,42	0,51
<b>MZ 4</b>	0,44	0,45	0,39	0,42	0,63

Tabelle 36: Varianzen der Prä-/Posterhebung Unterrichtsqualität der Gruppen 1-5

<b>Durchschnittliche Varianz über alle Skalen der Unterrichtsqualität</b>					
	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5
<b>Prä</b>	0,33	0,28	0,23	0,23	0,33
<b>Post</b>	0,31	0,21	0,26	0,30	0,38

Außerdem hatte die Lehrperson den Eindruck, dass das Interesse der Schülerinnen und Schüler an den Ergebnissen eher gering war und somit verkürzte sie die Vorstellungs- und Diskussionsphase der Schülerfeedbacks.

Auf die Frage, ob die Lehrperson in der Unterrichtsvorbereitung etwas änderte antwortete sie wie folgt:

*B Hm (.) kaum, kaum es ist so bei mir, dass ich manchmal Zeit habe dafür, viel vorzubereiten, manchmal muss es auch ganz schnell gehen und dann war das eigentlich jetzt nicht unbedingt jetzt dem Fragebogen geschuldet, der da kam. Andererseits versucht man natürlich schon überhaupt, was messbares zu machen, dann während des Unterrichts. Also man geht schon anders rein, finde ich und das ist gar nicht jetzt unbedingt die Vorbereitung, die man trifft. Denn oft sind auch die Stunden, die völlig unvorbereitet laufen ganz gut und andere, wo man sich stundenlang was überlegt hat, das geht dann plötzlich gar nicht, gibt's ja*

*auch. Aber ich glaube die Art und Weise wie man, wie man Auftritt ist vielleicht schon eine andere und da, ehm hat es sicherlich ein, die Befragung ein ein, Einfluss auf das [...].*

Die Lehrperson spricht also eher von einer anderen Einstellung, wie man in den Unterricht geht und weniger von einer intensivierte Vorbereitung.

#### **5.5.5.1 Charakterisierung Gruppe 5**

Zusammenfassend lässt sich die Gruppe 5 folgendermaßen beschreiben: Als einzige aller Gruppen weist sie negative Effekte der Unterrichtsqualität gemittelt über alle Skalen auf und dies obwohl die Werte zur Präerhebung schon die niedrigsten aller Gruppen darstellten. Diese Entwicklung spiegelt sich in den Schülerkurzfeedbacks nur zu MZ 1 und 2 wider. Im weiteren Verlauf der Intervention kommt es zu signifikanten positiven Entwicklungen. Dieser Widerspruch wird teilweise durch den Sonderstatus einer rückgemeldeten Schulstunde erklärt, deren Lerninhalt im Anschluss benotet wurde.

Insgesamt scheinen innerhalb der Schulklasse und auch zwischen Schulklasse und Lehrperson gesplante Verhältnisse und Konflikte vorzuliegen:

*B [...] natürlich jetzt grad mit der Klasse auch ganz andere Dinge noch verhandeln, es gab, es gab Fälle wo wo es um Ausgrenzung beziehungsweise Mobbing ging [...]*

Dies äußert sich auch in den, bereits in Abschnitt 5.4.1.7 vorgestellten, unsachlichen Kommentaren innerhalb des offenen Antwortformats der Schüler-Kurzfeedbacks. Ein weiterer Hinweis diesbezüglich ist in den Einschätzungen der sozial-/kognitiven Voraussetzungen der Schülerinnen und Schüler aus Lehrpersonensicht zu finden, die in jeder Skala den tiefsten Wert aller Gruppen einnehmen. Die Einschätzung der Schülerinnen und Schüler als kompetente Beurteiler von Unterricht zeigt den zweittiefsten Wert (Abschnitt 5.4.3 Tab. 33).

Die genannten Aspekte könnten bei der Lehrperson zu einer geringen Motivation geführt haben, den Unterricht umzustellen, beziehungsweise die Aussagen der Schülerinnen und Schüler nicht ernst zu nehmen, was wiederum die Schülerinnen und Schüler veranlasste, sich zu beklagen, dass sich zu wenig am Unterricht änderte (Tab. 31). Nichtsdestotrotz sahen gleichzeitig relativ viele Schülerinnen und Schüler eine Verbesserung der Disziplin im Laufe der Intervention (Tab. 31), was wiederum ein weiterer Hinweis auf das gesplante Verhältnis innerhalb des Klassenverbunds darstellt.

Durch das Verfahren des Kurzfeedbacks gelang es nicht die, aus pädagogisch/didaktischer Sicht, schwierige Konstellation aufzulösen. Hinsichtlich der bewerteten Unterrichtsqualität



scheint sich die Situation gar verschlechtert zu haben. Dies könnte damit erklärt werden, dass manche Schülerinnen und Schüler über die wenigen Veränderungen während der Intervention enttäuscht waren und sie den Unterricht deshalb abwerteten.

### 5.5.6 Zusammenfassung der Einzelanalysen

Vergleicht man die fünf Gruppen miteinander, lassen sich, in Bezug auf die Entwicklung der Unterrichtsqualität, Auffälligkeiten identifizieren, die zur vorläufigen Beantwortung der Forschungsfrage 3 führen:

*Lassen sich Muster identifizieren, die eine positive Wirkung von Schülerfeedback auf die Unterrichtsqualität begünstigen könnten? (FF 3)*

1. Je seltener eine Lehrperson die Schüleraussagen als falsch interpretierte, beziehungsweise die Kritikpunkte als unkontrollierbar attribuierte, desto positiver waren die Effektstärken.
2. Je häufiger und mannigfaltiger die Lehrperson den Unterricht während der Intervention veränderte, desto positiver waren die Effektstärken.
3. Lehrpersonen, die sich den Unterricht, aufgrund des Wissens, dass dieser rückgemeldet wird, nochmals in seiner Struktur bewusst machten, zeigten höhere Effektstärken.
4. Lehrpersonen, die das Beurteilungsvermögen der Schülerinnen und Schüler bezüglich Unterricht als niedrig, die Lernvoraussetzungen als ungünstig, das Schüler-Lehrpersonen-Verhältnis als weniger gut und die Klasse als leistungsmäßig heterogen einschätzten, wiesen geringere Effektstärken auf.
5. Bestanden bereits innerhalb des Klassenverbunds größere Konflikte und die Schülerinnen und Schüler nahmen zu wenig Verbesserungen während der Schülerfeedback-Intervention wahr, führte dies zu negativen Effekten.

Im folgenden Kapitel werden die Forschungsfragen und deren Antworten eingehend diskutiert.

## 6 Diskussion, Hypothesenbildung und Ausblick

Im Folgenden wird zunächst die methodische Konzeption der Studie diskutiert, daraufhin die Beantwortung der Forschungsfragen im Sinne einer Synthese kritisch reflektiert, um im Anschluss aus den gewonnenen Daten und Reflektionen empirisch gestützte Hypothesen zu bilden, die mit Hilfe einer quantitativ verwertbaren Stichprobengröße innerhalb einer Nachfolgestudie überprüft werden könnten. Diese bilden wiederum die Grundlage für den darauf folgenden *Ausblick*.

### 6.1 Instrumente und Design

Die kritische Diskussion der methodischen Konzeption fokussiert auf zwei Aspekte: die Prä-/Posterhebung und das neu entwickelte Schüler-Kurzfeedback-Verfahren (Intervention). Letzteres kann sowohl hinsichtlich seiner Praktikabilität als auch bezüglich seiner Gütekriterien beleuchtet werden.

Bevor auf die Einzelheiten der Instrumente und deren Güte eingegangen wird, ist zu erwähnen, dass die Teilnahme an der Studie auf Freiwilligkeit beruhte und somit eine Positiv-Selektion der Stichprobe vorliegen könnte. Die Daten sind somit vor dem Hintergrund dieser Problematik zu betrachten.

Insgesamt ergaben sich für alle Skalen der Schülerkurzfeedbacks hohe innere Konsistenzen, allerdings wäre kritisch anzumerken, dass die beiden Skalen der *Kognitiven Passung* eine Tendenz zum Deckeneffekt aufweisen. Eine entsprechende Anpassung der Items wäre demnach anzuraten.

Das eingesetzte Instrument Schüler-Kurzfeedback bezieht sich jeweils auf eine einzelne Unterrichtsstunde. Somit lassen sich zunächst keine Rückschlüsse auf die Tiefenstruktur und Qualität des Unterrichts im Allgemeinen ziehen. Wird es jedoch periodisch eingesetzt, scheint sich die Konstrukt-Validität hinsichtlich dieser Faktoren zu erhöhen. Es werden Tendenzen in den Daten erkennbar, die mit Tendenzen aus Messverfahren, die einen längeren Zeitraum betrachten, vergleichbar sind (QuaSSU-Fragebogen). So zeigt sich im Intergruppenvergleich, dass die Entwicklungen der Unterrichtsqualität aus der Prä-/Posterhebung in den meisten Gruppen (Ausnahme Gruppe 5) Ähnlichkeiten zu Entwicklungen der Schüler-Kurzfeedbacks aufweisen (je höher die Effekte des Prä-/Postvergleichs, desto mehr signifikante Steigungen in den Skalen der Kurz-Feedbacks über die vier Messzeitpunkte, vgl. Kap 5.4.1.6). Außerdem werden bei der isolierten Betrachtung einzelner Skalen ebenfalls z. T. ähnliche Entwicklungstendenzen zwischen vergleichbaren Skalen der beiden Erhebungsformen deutlich (z. B. *kognitiv aktivierende und strukturierende Elemente* vs. *Formal-kognitive Strukturiertheit* siehe Unterkapitel 6.2). In

einer quantitativ angelegten Folgestudie sollten dementsprechend weitere Überprüfungen bezüglich der Konstrukt-Validität des Schüler-Kurzfeedbacks erfolgen.

Da das Untersuchungsdesign der Prä-/Posterhebung auf viele Instrumente zurückgriff, die sich bereits in der Praxis bewährten, konnten diese größtenteils problemlos eingesetzt werden. Die Güte der Instrumente und Skalen haben sich in der vorliegenden Studie weitgehend bestätigt. Ausnahmen bildeten (1.) der Kurztest intrinsische Motivation (KIM), der, durch die Anpassung der Items, in einer Skala keine ausreichende Reliabilität aufwies und (2.) die schriftliche Befragung der Lehrpersonen bezüglich ihres konstruktivistischen Unterrichtshandelns.

Da das zweitgenannte Instrument auf Lehrpersonenseite eingesetzt wurde (kleine Stichprobe), versuchte man über konvergenzorientierte *Methoden-Triangulation* die Aussagekraft der Daten zu erhöhen, beziehungsweise die Zuverlässigkeit der Aussagen zu überprüfen. Hierzu wurden zunächst die Items, die das entsprechende Unterrichtshandeln beschreiben, in der Prä- und der Posterhebung eingesetzt. Die Lehrpersonen mussten in geschlossenen Items beschreiben, in welchem Maße sie diese Handlungen im Unterricht durchführen. Darüber hinaus mussten die Lehrpersonen in der Posterhebung zu jedem dieser Items angeben, ob sich die Häufigkeiten der jeweiligen Unterrichtshandlungen aufgrund der Feedbackphase verändert haben (siehe Abschnitt 5.2.6 und Anhang 8.1.5). Man hat also versucht, den Prä-/Postvergleich mit der direkt beschriebenen Handlungsänderung abzugleichen. Allerdings konnten zum einen keine signifikanten Differenzen im Prä-/Postvergleich festgestellt werden und zum anderen waren die Angaben aus der Triangulation in erheblichem Maße widersprüchlich. Dies kann vor dem Hintergrund einer bereits beschriebenen Kritik an der konvergenzorientierten *Methoden-Triangulation* gesehen werden, in der formuliert wird, dass

„nicht davon ausgegangen werden kann, dass jeweils der eine Ansatz das Gleiche zu Tage fördern wird wie der andere oder dass bei Diskrepanzen der Ergebnisse das eine (oder das andere) Resultat damit widerlegt sei. Vielmehr liegt schon im Verhältnis von Methode und Gegenstand bei der jeweiligen Einzelmethode eine solche Diskrepanz begründet, die es erforderlich macht, Kriterien zur Beurteilung von Übereinstimmungen und Diskrepanzen der Resultate unterschiedlicher Methoden zu entwickeln“ (Flick 2011, S. 17-18)

Bei einer quantitativen Nachfolgeuntersuchung ist eine erneute Überprüfung des Instruments hinsichtlich seiner Validität anzuraten.

Das Instrument KIM wurde ursprünglich für die intrinsische Motivation während des Besuchs außerschulischer Lernorte aus dem *Intrinsic Motivation Inventory* nach Deci & Ryan (2003) entwickelt (vgl. Wilde et al., 2009). Da die, für diese Studie umformulierte Version zur Messung der intrinsische Fachmotivation, nicht die erhofften Gütekriterien aufwies, wird empfohlen einen entsprechenden Kurztest neu zu entwickeln und auf seine Reliabilität zu überprüfen.

Eine Einschränkung muss auch bezüglich der Reliabilität der Skala *Inhaltliche Strukturiertheit* (Unterrichtsqualität aus Schülersicht, QuaSSU) gemacht werden, die zur Präerhebung einen Cronbachs  $\alpha$ -Wert von lediglich 0,55 aufwies. Die hochsignifikante positive Entwicklung von Prä zu Post und die teilweise hohen Effektstärken in den einzelnen Klassen sollten deshalb nicht

überbewertet werden. Wie es allerdings zur Diskrepanz zwischen den Cronbachs  $\alpha$ -Werten zwischen QuaSSU-Erhebung und Post-Erhebung auf der einen Seite und der Präerhebung auf der anderen Seite kam, wird aus den Daten nicht ersichtlich. Möglicherweise veränderte sich im Interventionszeitraum (auch aufgrund des Schüler-Kurzfeedback-Verfahrens) bei einigen Schülerinnen und Schülern die Interpretation der jeweiligen Items. Eine erneute Überprüfung dieser Skala und gegebenenfalls eine Neuformulierung der Items scheinen in diesem Falle angezeigt zu sein.

Bezüglich der leitfadengestützten Interviews erwies sich das Rahmenmodell der Unterrichtsentwicklung als hilfreich bei der Kategorienbildung. Die Neufokussierung innerhalb der *Rezeptionsphase* auf die jeweilige emotionale Ausprägung lohnte sich dahingehend, dass dadurch Hinweise auf eine negative Korrelation zwischen *sachlicher Rezeption* und *Motivation* aufgedeckt werden konnten.

Rückblickend hat sich innerhalb der methodischen Konzeption vor allem die komplementärorientierte Triangulations-Methode als nutzbringend erwiesen. Die quantitativen und qualitativen Daten aus *Unterrichtsqualität Prä-Post*, *Verlauf der Schüler-Kurzfeedbacks*, *Einschätzungen der sozial-kognitiven Voraussetzungen* und *der Beurteilungsfähigkeit von Unterricht* und die *Beschreibungen der Unterrichtsentwicklung* auf Schüler- und Lehrpersonenseite schufen in ihrer Gesamtbetrachtung ein differenziertes und typisierendes Bild verschiedener Einzelfälle und Gruppen, welche nicht nur deskriptiven Charakter aufweisen, sondern in Ansätzen auch erklärendes Potential besitzen. Die Entwicklung der Unterrichtsqualität als zentrales Ziel von Schülerfeedback zu formulieren, aufgrund dessen Gruppen und Einzelfälle zu bilden, diese daraufhin mit weiteren Daten aus der Studie zunächst zu untermauern und zum Schluss Ansätze für Erklärungen dieser Entwicklungen zu finden, erwies sich als fruchtbar. Das Verfahren wurde also dem komplementären Ziel der Triangulation gerecht, ein tieferes und weiteres Verständnis des Untersuchungsgegenstandes zu ermöglichen (Fielding & Fielding 1986, Flick 2011).

## **6.2 Lassen sich Hinweise identifizieren bezüglich des Einflusses von Schüler-Kurzfeedback auf a) ausgewählte Skalen der Unterrichtsqualität unter besonderer Berücksichtigung des Schüler-Lehrpersonen-Verhältnisses und b) die Unterrichtsentwicklung und das Unterrichtshandeln der Lehrperson? (FF 1)**

Die Beantwortung dieser Forschungsfrage bildet das zentrale Anliegen dieser Dissertation. Für die unter a) aufgeführte Unterrichtsqualität wurden zwei Datensätze herangezogen: zum einen der Verlauf der Schüler-Kurzfeedbacks und zum andern die Vergleiche der Prä-/Posterhebung über alle Schulklassen aus Sicht der Schülerinnen und Schüler.

Sieben der zwölf ausgesuchten Skalen wiesen im Durchschnitt über alle Schulklassen eine signifikante Verbesserung von Prä zu Post auf und keine Skala eine signifikant negative Entwicklung (es sei aber daran erinnert, dass Klassenweise auch negative Effekte auftraten). Vergleicht man diesbezüglich die Verläufe der Schüler-Kurzfeedbacks, findet man sowohl bestätigende als auch abweichende, aber nie widersprüchliche Tendenzen in den Daten. So zeigen beispielsweise die Daten der Skala *Disziplin* innerhalb der Schülerkurzfeedbacks positive Entwicklungen, während sich die Werte der vergleichbaren Skala *Klassenmanagement* der Prä-/Posterhebung nicht entwickelten. Allerdings findet die diesbezügliche Entwicklung der Schülerkurzfeedbacks Bestätigung in den Äußerungen der Schülerinnen und Schüler auf die Frage, welche Änderungen ihnen im Unterrichtsgeschehen aufgefallen sind. Hier wurde die Antwortkategorie *Verbesserung der Disziplin/Atmosphäre* mit 18 Mal am häufigsten genannt. Die Ursache für die Diskrepanz liegt wahrscheinlich in der, im Vergleich zu anderen Skalen, hohen Bewertung des Klassenmanagements in den acht Klassen der Gruppen 2 - 4 zur Präerhebung (Siehe Abschnitt 5.3.1 Tab. 27). Dies ist unter anderem mit der QULP-Studie vergleichbar, die ebenfalls gute bis sehr gute Werte für das Klassenmanagement hervorbrachte (vgl. Junge, Arx & Labudde 2014). Positive Entwicklungen, wie sie z. B. in Gruppe 1 (Klasse 10) zu sehen sind (siehe Tab. 26), konnten somit in den anderen Gruppen nur schwer abgebildet werden, negative Entwicklungen hingegen schon.

Die marginal signifikant positive Entwicklung der *kognitiv aktivierenden und strukturierenden Elemente* verläuft ähnlich wie die signifikant positiven Prä-/Postentwicklungen der vergleichbaren Skalen *Inhaltliche Strukturiertheit*, *Formal-kognitive Strukturiertheit* und *Interessanztheit*. Hier ist es außerdem möglich, dass die signifikant positive Steigung der Skala *Experimente als theoriegeleitetes Handeln* innerhalb der Kurzfeedbacks im Prä-/Postvergleich mit abgebildet wird, da die Items dieser Skala auch kognitiv strukturierende Elemente abfragen.

Die auftretenden Abweichungen in den Daten zwischen Schüler-Kurzfeedback und Prä-/Posterhebung sind jedoch nicht über zu bewerten. Erstens wurden zwar vergleichbare aber nicht die selben Items eingesetzt und zweitens handelte es sich bei den Schülerkurzfeedbacks um vier Momentaufnahmen einer (Doppel-) Stunde, während bei der Prä-/Posterhebung der Zeitraum vor, beziehungsweise während der gesamten Intervention betrachtet wurde (der bis zu zwölf (Doppel-) Stunden umfasste). Effekte, die durch eine intensivierte Vorbereitung aufgrund des Wissens, dass eine Feedback folgen wird (siehe Unterkapitel 5.5) oder zufällig vom Normalfall abweichende Stundenverläufe hätten somit die Daten der Schüler-Kurzfeedbacks stärker verzerren können.

Die durchschnittlichen Effekte des Prä-/Postvergleichs der Unterrichtsqualitäts-Skalen waren relativ gering. So zeigten nur vier Skalen einen durchschnittlichen Effekt von  $d > 0,2$ . Obwohl diese Ergebnisse aufgrund der kleinen Stichprobe nicht generalisiert werden können, fällt auf, dass die Ergebnisse in eine andere Richtung zeigen als die von Dittton & Arnold (2004), die eine

leicht signifikante Verschlechterung der Unterrichtsqualität nach einem einmaligen Schüler-Feedback feststellten. Dies kann als Hinweis gedeutet werden, dass ein häufigeres und kürzer getaktetes Schüler-Feedback die Wirkung auf die Unterrichtsqualität erhöht. Hingegen erinnern die Ergebnisse stark an die Untersuchungen im Hochschulbereich von Cohen (1980) und L'Hommedieu, Menges & Brinko (1990), die für Studentenfeedbacks ohne zusätzliche Beratung Effekte zwischen  $d = 0,18$  und  $d = 0,20$  ermittelten. Da die beiden Forschungsarbeiten wesentlich höhere Effekte bei einem Feedback mit zusätzlicher Beratung der Lehrperson ermittelten, liegt die Überlegung nahe, den Effekt eines zusätzlichen Coachings auch im Regelschulbereich zu untersuchen, zumal sich einige Lehrpersonen eine weitere Unterstützung im pädagogisch/didaktischen Umgang mit dem Schüler-Feedback wünschten (siehe auch Unterkapitel 6.4). Da die Effekte zwischen den Klassen auch ohne einen vergleichenden Einsatz eines Coachings z. T. stark variieren, ist davon auszugehen, dass weitere Faktoren die Wirkung von Schüler-Feedback beeinflussen. Diesen Phänomenen wird in Unterkapitel 6.4 nachgegangen.

Betrachtet man das *Schüler-Lehrpersonen-Verhältnis* als Aspekt der Unterrichtsqualität, so findet man auf Seiten der Schülerinnen und Schüler eine signifikante Verbesserung der Bewertungen und eine hohe zustimmende Antworthäufigkeit der entsprechenden Subkategorie innerhalb des Lehrpersoneninterviews. Allerdings ist innerhalb des geschlossenen Antwortformats der Lehrpersonen-Fragebögen keine Verbesserung festzustellen. Wahrscheinlich sind die mündlich beschriebenen Verbesserungen nicht groß genug, um sich in einer höheren Stufe der geschlossenen Items niederzuschlagen, beziehungsweise wäre bei bereits maximaler Bewertung zu Prä keine Steigerung mehr sichtbar zu machen, obwohl diese stattgefunden hat. Für diese Hypothese würde auch der sehr kleine Effekt von  $d = 0,14$  auf Schülerseite sprechen. Die Hypothese von Gärtner & Vogt (2013), dass Schülerfeedback die Schüler-Lehrpersonen-Beziehung begünstigt, kann aufgrund der geringen Effekte im vorliegenden Fall also nur in eingeschränktem Maße untermauert werden.

Bezüglich des Unterrichtshandelns und der Unterrichtsentwicklung lässt sich generell festhalten, dass die Lehrpersonen die Ergebnisse aus dem Feedback größtenteils nüchtern aufnahmen und meist den Schüleraussagen zustimmten. Davon wurde aber die Mehrheit der Aussagen als nicht veränderbar (unkontrollierbar) attribuiert. Sah man jedoch Möglichkeiten zur Veränderung (kontrollierbar attribuiert), bildeten sich vielfältige Änderungsabsichten, die wiederum meistens zu tatsächlichen Handlungsänderungen im Unterricht führten.

Allerdings kam es nicht nur durch die Rückmeldungen der Schülerinnen und Schüler zu Veränderungen im Unterrichtshandeln, sondern auch aufgrund des Wissens, dass ein Feedback auf die Stunde folgen wird (sechs Lehrpersonen machten entsprechende Angaben). Die Lehrpersonen reflektierten verstärkt über den noch nicht gehaltenen Unterricht, was folglich ihre Planung und Umsetzung beeinflusste. Dies steht stark im Einklang mit den Befunden von Gärtner & Vogt (2013, S. 265), die aufgrund dessen formulierten, dass „die Items eines solchen Befra-

gungsangebotes als Kommunikationsmittel von Indikatoren guten Unterrichts fungieren.“ Dabei betraf die Reflektion häufig die Struktur der Unterrichtsstunde (was sich in den oben beschriebenen Entwicklungen der entsprechenden Skalenwerte widerspiegelt) und den Umgang mit Experimenten. Letzteres wird nicht nur aus den Antworthäufigkeiten der Interviews abgeleitet (es berichteten sechs Lehrpersonen von einer stärkeren Thematisierung von Experimenten und dem wissenschaftliche Arbeiten) sondern auch untermauert durch die signifikant positive Steigung der Regressionsgeraden der Skala *Experimentieren als theoriegeleitetes Handeln* innerhalb der Schüler-Kurzfeedbacks.

Gemittelt über alle Schulklassen lässt sich insgesamt festhalten, dass sich sowohl in der Prä-/Posterhebung als auch im Verlauf der Schüler-Kurzfeedbacks viele Skalen signifikant positiv entwickeln (allerdings wurden innerhalb der Prä-/Posterhebung kleine Effektstärken gemessen), während keine Skala eine signifikant negative Entwicklung nimmt. Dies spricht für ein hohes Wirkungs-Potenzial des Schüler-Kurzfeedback-Verfahrens auf die Unterrichtsqualität.

Bezogen auf die Forschungsfrage und die gewonnenen empirischen Daten können demnach folgende begründeten Hypothesen formuliert werden:

*Hypothese 1:*

*Das eingesetzte Schüler-Kurzfeedback-Verfahren trägt im Durchschnitt zu einer Verbesserung der strukturierenden Elemente von Unterricht bei.*

*Hypothese 2:*

*Das eingesetzte Schüler-Kurzfeedback-Verfahren trägt im Durchschnitt zu einer Verbesserung der kognitiv aktivierenden Elemente von Unterricht bei.*

*Hypothese 3:*

*Das eingesetzte Schüler-Kurzfeedback-Verfahren verbessert im Durchschnitt das Schüler-Lehrpersonen-Verhältnis.*

*Hypothese 4:*

*Das eingesetzte Schüler-Kurzfeedback-Verfahren verstärkt im Durchschnitt auf Schülerseite das theoriegeleitete Handeln während des Experimentierens.*

*Hypothese 5:*

*Das eingesetzte Schüler-Kurzfeedback-Verfahren führt auf Lehrpersonenseite während der Unterrichtsentwicklung zu einer intensivierten Auseinandersetzung mit strukturierenden Aspekten von Unterricht.*

*Hypothese 6:*

*Das eingesetzte Schüler-Kurzfeedback-Verfahren führt auf Lehrpersonenseite während der Unterrichtsentwicklung zu einer intensivierten Auseinandersetzung mit der theoretischen Einbettung von Experimenten im Unterricht.*

### **6.3 Welche Entwicklungen nehmen Schülerinnen und Schüler im Unterricht innerhalb des Schüler-Kurzfeedback-Verfahrens wahr? (FF2)**

Wie im vorigen Unterkapitel bereits ausgeführt, bewerten die Schülerinnen und Schüler einige Skalen der Unterrichtsqualität nach der Intervention besser als davor. In diesem Abschnitt steht die Frage im Zentrum, ob die Schülerinnen und Schüler die Veränderungen bewusst wahrnehmen, und wenn ja, wie diese Veränderungen aussahen.

Da davon ausgegangen werden kann, dass die Schülerinnen und Schüler mit einer gewissen Erwartungshaltung in solch ein Schüler-Feedback-Verfahren gehen, nämlich, dass ihre Rückmeldungen zu Veränderungen im Unterrichtshandeln der Lehrperson führt, sollten die Schülerinnen und Schüler im besten Falle möglichst viele Veränderungen klar erkennen können. Ist dies nicht der Fall und die Erwartungen werden enttäuscht, könnte dies eventuell zu einer Abwertung der Unterrichtsqualität führen. Um einen solchen Effekt zu verhindern, sollten die Schülerinnen und Schüler zunächst den Eindruck gewinnen, dass die Lehrperson das Verfahren ernst nimmt. Dem stimmten die Schülerinnen und Schüler im Durchschnitt zu (Mittelwert = 3,44). Allerdings schlägt sich dies, aus deren Sicht, nur wenig in bewusst wahrgenommener Veränderung des Unterrichts nieder (Mittelwert = 2,15). Falls die Schülerinnen und Schüler Veränderungen wahrgenommen haben, beschrieben sie meistens eine verbesserte Disziplin/Atmosphäre. Manche berichteten davon, dass die Lehrperson besser erklärt hat und dass mehr Experimente durchgeführt wurden.

Diese Veränderungen wurden größtenteils auf das Feedback-Verfahren zurückgeführt (Wert = 3,14). Diese Befunde stehen einerseits im Widerspruch zu den Ergebnissen aus der Prä-/Posterhebung dieser Untersuchung und andererseits auch zu den Ergebnissen von Gediga (2000), die im Hochschulbereich eine gegenteilige Tendenz feststellte, nämlich kaum Effekte in den geschlossenen Items der Prä-/Posterhebung, dafür wurde bei direkter Befragung der Studenten hinsichtlich wahrgenommener Veränderungen, häufig von solchen berichtet. Der erste Widerspruch wurde teilweise bereits im Ergebnisteil erklärt, indem aufgezeigt wurde, dass die bewusst wahrgenommenen Veränderungen mit den Differenzen der Unterrichtsqualitäts-Skalen sehr häufig positiv korrelieren (siehe Abschnitt 5.2.4.2). Das heißt, die bewusste Wahrnehmung von Unterrichts-Veränderung eines Lernenden erhöht die Wahrscheinlichkeit, dass dieser den



Unterricht zur Posterhebung besser beurteilt als zur Präerhebung. Des Weiteren ließe sich der Widerspruch mit der mangelnden Fähigkeit vieler Schülerinnen und Schüler erklären, Unterricht auf didaktischer/pädagogischer Ebene adäquat zu analysieren. Dies entspricht auch der Beobachtung von Clausen (2002), der bei differenzierter neutraler Beschreibung und Bewertung des unterrichtlichen Geschehens, außenstehenden Beobachter gegenüber den Schülerinnen und Schülern bevorzugt.

Pointiert ausgedrückt: Die Schülerinnen und Schüler beschreiben eine Verbesserung des Unterrichts, eine konkrete Veränderung nehmen sie jedoch eher nicht wahr; ähnlich der mangelnden Fähigkeit eines Menschen die Bewegung eines Stundenzeigers wahrzunehmen, er im Nachhinein aber eine andere Zeit feststellt. Dies wäre auch eine Erklärung für die widersprüchlichen Ergebnisse bezogen auf die Studie im Hochschulbereich von Gediga (2000): Studierende sind aufgrund des höheren Alters und der höheren Bildung möglicherweise besser in der Lage, differenzierte Veränderungen im Unterricht wahrzunehmen als Schülerinnen und Schüler.

Wie aus den Interviews mit den Lehrpersonen hervorgeht, nehmen diese viele Änderungen im Unterrichtshandeln vor, die allerdings von den Schülerinnen und Schülern nur selten wahrgenommen werden. Demnach wäre den Lehrpersonen angeraten, den Schülerinnen und Schülern diese Änderungen während des Unterrichts zu verdeutlichen. Die Gefahr des Nichterfüllens einer erhöhten Erwartungshaltung würde sich dadurch eventuell verringern, was sich wiederum auf das Schüler-Lehrpersonen-Verhältnis und die Unterrichtsqualität positiv auswirken könnte.

Zusammengefasst ergeben sich folgende Hypothesen:

*Hypothese 7:*

*Schülerinnen und Schüler nehmen vereinzelt Veränderungen im Unterrichtshandeln im Rahmen des Schüler-Kurzfeedback-Verfahrens wahr.*

*Hypothese 8:*

*Die wahrgenommenen Veränderungen im Unterrichtshandeln im Rahmen des Schüler-Kurzfeedback-Verfahrens erhöhen sich, wenn die Lehrperson das veränderte Unterrichtshandeln den Schülerinnen und Schülern transparent macht.*

Weiterführende Hypothesen:

*Hypothese 9:*

*Das Transparentmachen der Veränderungen im Unterrichtshandeln aufgrund des Schülerfeedbacks erhöht die positive Wirkung des Schüler-Kurzfeedback-Verfahrens auf das Schüler-Lehrpersonen-Verhältnis.*

*Hypothese 10:*

*Das Transparentmachen der Veränderungen aufgrund des Schülerfeedbacks erhöht die positive Wirkung des Schüler-Kurzfeedback-Verfahrens auf die Unterrichtsqualität.*

#### **6.4 Lassen sich Muster identifizieren, die eine positive Wirkung von Schülerfeedback auf die Unterrichtsqualität begünstigen könnten? (FF 3)**

Die bisher untersuchten und diskutierten Fragestellungen dienten als Ausgangspunkt für eine differenzierte Betrachtung der möglichen Ursachen für eine positive Entwicklung der Unterrichtsqualitäts-Faktoren. Dies ist auch deshalb von hoher Relevanz, weil die Unterschiede der durchschnittlichen Effekte zwischen den Klassen hoch sind. Am deutlichsten wird dies im Vergleich zwischen Klasse 10 ( $d = 0,57$ ) und Klasse 3 ( $d = -0,34$ ). Die Heterogenität der auftretenden Effekte deckt sich auch mit Untersuchungen aus dem Hochschulbereich (vgl. Rindermann 2009). Es ist also für den Einsatz von Schüler-Feedback entscheidend zu wissen, welche Maßnahmen dieses fruchtbar macht beziehungsweise, was wahrscheinlich noch wichtiger ist, die Ursachen für einen negativen Impact eines solchen Verfahrens zu kennen. Mit Hilfe der *komplementärorientierten Triangulation* konnten einige Hinweise diesbezüglich ermittelt werden. Hierfür wurde auf Grundlage der Unterrichtsqualitäts-Daten der Prä-/Posterhebung eine Clusteranalyse durchgeführt, was zur Bildung von fünf Gruppen führte.

Auf diese Weise konnten verschiedene Muster identifiziert werden, die wiederum Hinweise zu den Ursachen unterschiedlicher Entwicklungen der Unterrichtsqualitätsfaktoren liefern. So zeigt sich bei der Einschätzung der *kognitiven und sozialen Voraussetzungen* und des *Beurteilungsvermögens der Schülerinnen und Schüler von Unterricht*, dass Lehrpersonen der Gruppen, die keine beziehungsweise negative Effekte zeigten, den Schülerinnen und Schülern im Durchschnitt schlechtere *Lernvoraussetzungen* zuschrieben und die Klassen als *leistungsheterogener* einstufen. Außerdem schätzten sie das *Schüler-Lehrpersonen-Verhältnis* und das eigene *Klassenmanagement* vergleichsweise schlechter ein. Des Weiteren hielten sie, im Vergleich zu den anderen Schulklassen, Schülerinnen und Schüler für weniger gute Beurteiler von Unterricht.

Der letzte Punkt findet seine Fortführung in der Rezeption der Rückmeldungen, wie sie in den Lehrpersoneninterviews geäußert wurden. So antworteten die Lehrpersonen dieser beiden Gruppen am häufigsten, dass die Schüleraussagen falsch oder nicht veränderbar (sie attribuieren unkontrollierbar/stabil) seien, was sich nach Strijbos & Müller (2014) negativ auf die weitere mögliche Handlungsänderungen auswirkt. Dies wurde vor allem bei der Gruppe 5 (Klasse 10) im weiteren Verlauf der Interviews bestätigt. Die Lehrperson dieser Klasse beschrieb die we-

nigsten Änderungsabsichten und daraufhin auch die wenigsten tatsächlich durchgeführten Handlungsänderungen.

Diese Ergebnisse finden in den Schülerkommentaren zu dem offenen Item *So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich schlecht, weil...* Bestätigung. Gruppe 5 weist in den beiden Kategorien *zu wenig geändert* und *nichts geändert* zusammengekommen eine ca. dreimal so hohe Antworthäufigkeit pro Schüler/-in auf wie Gruppe 4 mit dem zweithöchsten Wert.

Die Gruppe 1 mit den höchsten Effekten der Unterrichtsqualität im Prä-/Postvergleich zeigte diesbezüglich ein gegenteiliges Bild. Hinsichtlich der Subkategorie *Die Schüleraussage ist richtig und veränderbar* war die Antworthäufigkeit am höchsten (es wurde kontrollierbar/variabel attribuiert) und ebenso wurden Änderungsabsichten und tatsächliche Handlungsänderungen am häufigsten beschrieben. Außerdem berichtete in dieser Gruppe kein/e Schüler/-in, dass sich zu wenig oder nichts am Unterricht geändert hat. Über alle Gruppen hinweg zeigt sich in den Interviews die Tendenz, dass mit steigender Effektstärke nicht nur die Häufigkeit sondern auch die Mannigfaltigkeit der berichteten Veränderungen steigt.

Wie bereits in Unterkapitel 6.1 beschrieben, setzte allein die Erwartung, dass ein Schüler-Feedback folgt, Reflexionen über den Unterricht in Gang, die sich vor allem auf dessen Struktur und auf den Einsatz von Experimenten und deren theoretischer Einbettung bezogen. Interessant ist hierbei, dass dies eher in Gruppen geschah, die eine größere Effektstärke in der Unterrichtsqualitätsentwicklung aufweisen.

Zusammengefasst scheinen drei Aspekte für den Erfolg des Schüler-Feedback-Verfahrens besonders entscheidend zu sein:

1. Ein hohes Vertrauen in die Richtigkeit der Schülerrückmeldung als Grundlage für eine handlungsfördernde Attribution im Unterrichtsentwicklungsprozess. Dieses Vertrauen und die jeweiligen Attributionsstile beruhen wahrscheinlich sowohl auf einer persönlichen Grundhaltung gegenüber der Güte von Schülerurteilen als auch auf einer differenzierten Einschätzung der Schulklasse. So zeigt sich bei genauerer Betrachtung der Gruppe 5, dass die unterdurchschnittliche *Einschätzung der Beurteilungsfähigkeit* auf Lehrpersonenseite von den Schülerinnen und Schülern der Klasse teilweise bestätigt wird. Denn diese äußerten am häufigsten, dass die Mitschüler/-innen das Feedback nicht ernst nähmen. Ein weiteres Indiz hierfür sind die relativ vielen unsachlichen Kommentare innerhalb der offenen Antworten des Schüler-Kurzfeedbacks. Ungünstige Konstellationen im Klassenverbund können also dem Erfolg eines Schüler-Feedbacks entgegenstehen.
2. Die Häufigkeit und Mannigfaltigkeit der Handlungsänderungen im Rahmen der Unterrichtsentwicklung. Dies wiederum ist zunächst davon abhängig, ob viele und unterschiedliche Kritikpunkte von den Schülerinnen und Schülern geäußert und daraufhin

aufgegriffen werden können. Eine weitere Grundvoraussetzung für eine mannigfaltige Handlungsänderung bildet das bereits unter dem ersten Aspekt beschriebene Vertrauen in das Schülerurteil.

3. Inwieweit die Lehrperson, aufgrund des zu erwartenden Schüler-Feedbacks, den eigenen Unterricht reflektiert und die Unterrichtsentwicklung in differenzierter Form anpasst.

Neben diesen drei Aspekten könnte ein vierter eine wichtige Rolle für den Erfolg eines Schüler-Feedbacks spielen. Dies geht vor allem aus Äußerungen der Lehrpersonen während der Interviews sowie auch aus Studien im Hochschulbereich hervor. Es scheint ein großes Bedürfnis auf Lehrpersonenseite zu bestehen, weitere Hinweise von Expertenseite im Umgang mit den Daten der Schüler-Kurzfeedbacks zu erhalten. Studien aus dem Hochschulbereich berichten diesbezüglich von einer deutlich verbesserten Wirkung auf die Lehrqualität beim Einsatz einer solchen Maßnahme (Coaching, Abschnitt 2.2.1). Es ist möglich, dass gerade in Schulklassen, in denen ein negativer Effekt durch das Verfahren droht, dieser eventuell verhindert werden kann. Dabei ist jedoch zu beachten, dass dadurch die angestrebte Einfachheit des Verfahrens stark eingeschränkt wird. Trotzdem sollte diesem Aspekt in Zukunft, vor allem im Hinblick auf die praktische Lehrpersonenaus- und Weiterbildung, verstärkt Beachtung geschenkt werden. Da die praktische Ausbildung ohnehin von einem Mentor/Coach begleitet wird, könnte das Schüler-Kurzfeedback-Verfahren als Perspektivenerweiterung, vor allem im Hinblick auf die Sichtbarmachung der Tiefenstrukturen, begleitend eingesetzt werden. Geht man davon aus, dass die Tiefenstrukturen ein entscheidendes Element der Unterrichtsqualität darstellen, wären die Ergebnisse aus den Schüler-Kurzfeedbacks ein wichtiger Baustein im Prozess des Unterrichten Lernens.

Diese vier Aspekte münden in folgende Hypothesen:

*Hypothese 11:*

*In Schulklassen, deren Lehrpersonen einer Schülerrückmeldung eine hohe Validität zuschreiben, ist der Effekt des Schüler-Kurzfeedback-Verfahrens auf die Unterrichtsqualität höher.*

*Hypothese 12:*

*In Schulklassen, deren Lehrpersonen vergleichsweise häufiger Änderungsmaßnahmen durchführen, ist der Effekt des Schüler-Kurzfeedback-Verfahrens auf die Unterrichtsqualität höher.*

*Hypothese 13:*

*In Schulklassen, deren Lehrpersonen den Unterricht in Erwartung eines Schüler-Feedbacks anpassen, ist die Wirkung des Schüler-Kurzfeedback-Verfahrens auf die Unterrichtsqualität höher.*

Weiterführende Hypothese:

*Hypothese 14:*

*Ein Schüler-Feedback, das mit Coaching-Maßnahmen kombiniert wird, erhöht den Effekt des Schüler-Kurzfeedbacks auf die Unterrichtsqualität.*

Möchte man aufgrund der Erkenntnisse Maßnahmen ergreifen, die das Schüler-Feedback-Verfahren auf einfache Weise in seiner Wirkung erfolgreicher machen sollen, so ist wohl zentral, das Vertrauen der Lehrpersonen in die Validität der Schülerrückmeldungen zu stärken, da dies die weiteren Unterrichtsentwicklungsprozesse positiv beeinflusst.

## **6.5 Wie schätzen Lehrpersonen und Schülerinnen und Schüler die Praktikabilität des Schüler-Kurzfeedback-Verfahrens ein? (FF 4)**

Neben der Frage zur Wirkung des Schüler-Kurzfeedbackverfahrens war von Bedeutung, ob dieses sowohl auf Lehrpersonen als auch auf Schülerinnen- und Schüler-Seite als praktikabel eingeschätzt wird. Nur so besteht die Chance, dass, im Falle einer nachgewiesenen Wirkung, die Forschungsergebnisse für die Unterrichtspraxis fruchtbar gemacht werden können.

Auf Lehrpersonenseite wurde das Verfahren insgesamt als praktikabel eingeschätzt, vor allem weil es zeitökonomisch einsetzbar ist, die Schülerinnen und Schüler ernst genommen werden und die Kommunikation zwischen Lehrperson und Schülerinnen und Schüler anregt.

Kritisch wird jedoch angemerkt, dass durch letztgenannten Punkt die als wichtig erachtete Anonymität aus den Fragebögen teilweise verloren geht. Demgegenüber wird jedoch positiv gesehen, dass sowohl die Lehrpersonen als auch die Schülerinnen und Schüler dazu angeregt wurden, genauer über den Unterricht nachzudenken. Dabei erachtete man die Skalen der Schüler-Kurzfeedback-Bögen als hilfreich und sinnvoll, weil sie einen Anhaltspunkt für die Diskussion und die Unterrichtsverbesserungen darstellen.

Die Konzentration auf Skalen, welche die Tiefenstruktur von Unterricht abbilden, scheint demnach nicht nur ein zentraler Aspekt aus Wissenschaftsperspektive zu sein, um die Unterrichtsqualität zu messen; sie erwies sich auch in der Schulpraxis als Hilfreich, um Entwicklungen im Unterrichtshandeln anzuregen. Hinsichtlich dessen gaben drei Lehrpersonen an, dass die Fragebogen-Ergebnisse für die Unterrichtsentwicklung entscheidender waren als die Diskussionen

darüber. Trotzdem erachten die Lehrpersonen die Vorstellung der Ergebnisse als gut und wichtig, weil sie glauben, dass die Schülerinnen und Schüler das Bedürfnis haben, die Ergebnisse zu erfahren und die Wertschätzung der Schülermeinung verdeutlicht wird. Des Weiteren wurde bezüglich der Diskussion eine Perspektivenerweiterung auf beiden Seiten gesehen: Die Schülerinnen und Schüler sehen, wie ihre eigene Rückmeldung im Vergleich zur Klasse aussieht und unter Umständen auch, wie ihr eigenes Verhalten bei den Mitschülern ankommt; und die Lehrperson bekommt zu den eher abstrakten Ergebnissen aus den Fragebögen konkrete Anhaltspunkte zur Verbesserung des Unterrichts.

Dies verdeutlicht die Sichtbarmachung der von Clausen (2002) beschriebenen *perspektiven-spezifischen Validität* (siehe Abschnitt 2.1.2). Und zwar sowohl zwischen Lehrpersonen- und Schülerseite als auch zwischen den Schülerinnen und Schülern innerhalb der Klasse. Diese Sichtbarmachung kann nicht nur dazu genutzt werden, das gegenseitige Verständnis von entsprechenden (Unterrichts-)Handlungen zu fördern oder die Unterrichtsentwicklung entsprechend anzupassen, sondern auch alle Beteiligten für das Phänomen zu sensibilisieren, dass ein und der selbe Sachverhalt aus unterschiedlichen Perspektiven verschieden beschrieben werden kann.

Kritisch wurde bei einigen Lehrpersonen und Schülerinnen und Schülern der hohe Zeitaufwand gesehen (wobei anzumerken ist, dass einige Lehrpersonen das genaue Gegenteil formulierten). Die Kürzung des Fragebogens um eine Skala würde sowohl die Zeit des Ausfüllens als auch der Vorstellung und Diskussion beschleunigen. Möglich wäre dies durch den Zusammenzug der beiden Skalen der kognitiven Passung zu einer Skala, in welcher der mittlere Wert den Optimalwert darstellt. Hierfür müssten allerdings die Items neu formuliert und die Skala einer erneuten Pilotierung unterzogen werden.

Da manche Lehrpersonen über den Zeitraum der Intervention einen Motivationsverlust bezüglich des Ausfüllens der Fragebogen bei einigen Schülerinnen und Schüler wahrnahmen, sollte die Verfahrensdauer und Rückmeldehäufigkeit nicht erhöht werden. Aber auch eine Verringerung ist nicht angezeigt, da erstens das Phänomen des Motivationsverlusts nicht flächendeckend auftrat und zweitens aufschlussreiche Daten verloren gingen. Ein möglichst optimaler Einsatz des Verfahrens bezüglich Intervall und Dauer müsste in weiteren Studien ermittelt werden, beziehungsweise sollte in der Praxis nach individuellen Bedürfnissen und Einschätzungen der Lehrpersonen entsprechend variiert werden.

Obwohl den Lehrpersonen Hinweise zur Interpretation der aufbereiteten Rückmeldungen gegeben wurden, taten sich manche Lehrpersonen diesbezüglich schwer. Hier ist zu überprüfen, ob statt eines Säulendiagramms eine Punktwolke aller Schülerinnen und Schüler zuzüglich gekennzeichnetem Mittelwert und Standardabweichung, die Interpretierbarkeit der Daten verbessert. Auch wäre zu diskutieren, statt einer 4-Stufigen, eine 5- oder 6-Stufige Likert-Skala zu verwenden, um auftretende Unterschiede zwischen den einzelnen Messzeitpunkten möglicher-

weise besser sichtbar zu machen. Diese Maßnahmen würden sich wahrscheinlich auch positiv auf die Verständlichkeit der Ergebnisdarstellung auf Schülerinnen- und Schüler-Seite auswirken.

Zwei Lehrpersonen äußerten das Bedürfnis, pädagogische/didaktische Handlungsanweisungen zu auftretenden Rückmeldeergebnissen zu erhalten. Dies scheint nur im Rahmen eines, bereits oben erwähnten, Coachings möglich zu sein, da dies eine genaue Diagnose der Ursachen benötigt. Im Rahmen eines solchen Coachings könnte auch einer weiteren Forderung nachgekommen werden, nämlich Handlungstipps zu geben, wie man die Schülerinnen und Schüler während der Ergebnisdiskussion besser miteinbezieht und sie zu "Beteiligten" macht.

Manche Lehrpersonen vermissten außerdem Vergleichswerte zu den eigenen Daten, um diese in ihrer Ausprägung besser einschätzen zu können. Es ließe sich überlegen, ob in zukünftigen Studien oder beim Einsatz des Verfahrens in elektronischer Form in der Praxis, solche Daten zur Verfügung gestellt werden sollten. Das Für und Wider einer solchen Maßnahme soll hier jedoch nicht vertiefend diskutiert werden.

Insgesamt erfuhr das Schüler-Kurzfeedback-Verfahren einen großen Zuspruch auf Lehrpersonenseite. Dies wird dadurch deutlich, dass keine Lehrperson im Laufe der Studie aus dem Verfahren ausstieg und vor allem, dass neun Lehrpersonen das Verfahren in der Praxis einsetzen würden, wenn es in elektronischer Form vorläge und somit die Daten automatisch ausgewertet würden.

Auf Seiten der Schülerinnen und Schüler wird das Feedback-Verfahren wesentlich häufiger positiv als negativ bewertet. Dabei wird vor allem hervorgehoben, dass sie in die Lage versetzt werden, anonym und differenziert ihre Meinung zum Unterricht kundzutun. Die Schülerinnen und Schüler haben also ein großes Interesse sich zum Unterricht (sei es positiv oder negativ) zu äußern und gehört zu werden. Sie werden dadurch vom Konsumenten und Empfänger von Unterricht in gewissem Maße zum Mitgestalter. Hierdurch könnte das Verantwortungsgefühl für gelingenden Unterricht auf Schülerseite steigen. Als Indiz hierfür kann die Aussage einer Lehrperson gesehen werden, die berichtete, dass sie ergriffene Veränderungen/Maßnahmen leichter legitimieren konnte.

Dass die Unterschiede im Verantwortungsgefühl zwischen den Schüler/-innen wahrscheinlich groß sind, zeigen zum einen die Häufung sachferner Kommentare in einigen Schulklassen im Verlauf des Verfahrens und zum anderen die bereits oben erwähnten Aussagen mancher Schülerinnen und Schüler (vor allem in Gruppe 5), die der Auffassung waren, dass manche Mitschüler die Feedbacks nicht ernst genug nehmen. Dies sagt sowohl etwas über die Einstellung der beurteilten Mitschüler aus als auch über die Schülerinnen und Schüler, die diese Aussage tätigten. Diese würden wahrscheinlich gerne Verantwortung übernehmen und Mitgestalten, befürchten aber, dass dies durch das Verhalten mancher Mitschüler nicht möglich ist.

Um das Verantwortungsgefühl bei den entsprechenden Schülerinnen und Schülern zu steigern beziehungsweise ein Absinken aufzuhalten, wäre zu empfehlen, dieses den Schülerinnen und Schülern im Klassengespräch zu verdeutlichen und die unsachlichen Kommentare während der Ergebnisvorstellung nicht zu veröffentlichen, sodass die Mitschüler nicht angeregt werden, ebenfalls unsachliche (humoristische) Kommentare zu verfassen. Auch hier könnte ein Coaching helfen, die Lehrperson für entsprechende Prozesse zu sensibilisieren.

Zusammenfassend lässt sich festhalten, dass sowohl die Schülerinnen und Schüler als auch die Lehrpersonen das Schüler-Kurzfeedback-Verfahren, trotz einiger Kritikpunkte, als praktikabel einstufen. Bezüglich der Kritikpunkte wurden Optimierungsvorschläge unterbreitet. Einzelne daraus (Coaching, Zusammenfassung der beiden Skalen *Kognitive Passung* zu einer Skala, Verfügbarmachung von Vergleichsdaten usw.) sollten für eine Nachfolgestudie oder für eine Umsetzung in der Praxis ernsthaft in Erwägung gezogen werden.

Aufgrund der Mannigfaltigkeit der Beschreibungen und der Intention, dass die Beantwortung dieser Forschungsfrage auch dazu dienen soll, den Einsatz des Verfahrens in der Schul- und Forschungspraxis zu verbessern, wird darauf verzichtet, Hypothesen zu bilden, um die argumentative Breite der Aussagen nicht einzuschränken.

## **6.6 Weitere relevante Aspekte des Schüler-Kurzfeedback-Verfahrens**

Aus der Studie stechen drei Punkte hervor, die zwar nicht in direktem Zusammenhang mit den Forschungsfragen stehen, aber für einen Einsatz des Schüler-Kurzfeedback-Verfahrens in Wissenschaft und Praxis von Bedeutung sind:

1. Es stellt sich die Frage, ob der Einsatz eines solchen Verfahrens in einer Klasse, in der bereits erhebliche soziale Konflikte vorliegen, sinnvoll ist. Die Studie lieferte Hinweise, dass in solchen Klassen im Einzelfall eine Verschlechterung der beurteilten Unterrichtsqualität durch das Verfahren hervorgerufen werden kann. Es scheint für den Einsatz eines Schüler-Feedback-Verfahrens wichtig zu sein, dass die erhobenen Daten bezüglich der didaktischen Zielintentionen nicht von sozialen Konflikten im Klassenverbund überlagert werden. Für solche Fälle ist zu prüfen, inwiefern ein regelmäßiges Schüler-Feedback nur im Zusammenhang mit einem gezielten Coaching sinnvoll ist.
2. Wie nachhaltig ist das Feedbackverfahren? Die Lehrperson, welche die höchsten Effekte in der Unterrichtsqualitätsentwicklung erzielte, berichtete, dass sie während des Verfahrens einen ungewöhnlich hohen Aufwand betrieb und dass dieser in der alltäglichen Unterrichtspraxis nicht aufrecht zu erhalten sei. Auch hier könnten durch coachende Begleitung, sei dies durch einen externen Experten oder einen Kollegen, die relevanten Aspekte, z. B. zur Verbesserung der Tiefenstruktur, hervorgehoben werden, sodass durch diese Fokussierung die Veränderungen im Unterrichtshandeln eingegrenzt und



somit der Aufwand verringert würde. Außerdem ist zu empfehlen, die Nachhaltigkeit des Verfahrens in weiteren Studien (Follow-up-Design) zu untersuchen.

3. Die Skala *Experimente als theoriegeleitetes Handeln* zeigte gemittelt über alle Schulklassen im Verlauf der Intervention eine signifikant positive Entwicklung. Deshalb ist zu vermuten, dass neben den bereits untersuchten Skalen auch Entwicklungen im theoriegeleiteten Einsatz von Experimenten im Prä-/Postvergleich zu erwarten sind. Hierfür sollten entsprechende Items entwickelt und eingesetzt werden.

## **6.7 Rückblick und Ausblick**

Das zentrale Ziel dieser Untersuchung lag im Auffinden von Hinweisen zur Wirkung und Wirkmechanismen eines neu entwickelten Schüler-Kurzfeedback-Verfahrens auf die Unterrichtsentwicklung und Unterrichtsqualität. Es zeigte sich, dass die Lehrpersonen aufgrund des Verfahrens ihr didaktisches und pädagogisches Unterrichtshandeln auf unterschiedliche Weise veränderten, sei dies beispielsweise durch einen stärkeren Theoriebezug der eingesetzten Experimente, eine Verbesserung der Unterrichtsdisziplin oder eine kleinschrittigere Strukturierung des Unterrichts. Außerdem wurde in vielen Schulklassen eine Verbesserung von Unterrichtsqualitätsskalen festgestellt.

### **6.7.1 Folgerungen für die Naturwissenschaftsdidaktik**

Die vorgestellten Ergebnisse mündeten in zahlreichen Hypothesen, deren generelle Gültigkeit überprüft werden sollten. Zu diesem Zweck könnten quantitative Folgestudien durchgeführt werden, in welchen das Verfahren mit ausreichend großen Stichproben kontrollgruppenbasiert untersucht werden kann. Zudem wäre ein Follow-up-Design anzustreben, um die Nachhaltigkeit der zu erwartenden Effekte zu überprüfen. In weiterführenden Studien könnte außerdem eine stärkere Fokussierung auf den Lerneffekt durch Schüler-Kurzfeedback gelegt werden. Diesbezüglich wäre denkbar, das Konzeptverständnis oder die Kenntnisse über naturwissenschaftliche Arbeitsweisen in einem Prä-/Postvergleich zu erheben. Weiterführende Forschungsfragen könnten lauten:

- Inwieweit trägt ein Schüler-Kurzfeedback-Verfahren zu einer Verbesserung des Konzeptverständnisses im Fach Physik bei?
- Verbessert ein Schüler-Kurzfeedback-Verfahren die Kenntnisse über naturwissenschaftliche Arbeitsweisen?

Innerhalb solcher Studien wäre es begrüßenswert, wenn die, in der vorliegenden Arbeit gefundenen Hinweise zu Mustern oder Ansatzpunkten, die besonders positive Effekte begünstigen,

durch entsprechende Maßnahmen verstärkt und vergleichend eingesetzt würden. Hierzu sollte beispielsweise eine Aufklärung über die hohe Güte der Schülerurteile gehören, um das Vertrauen der Lehrpersonen in die Schülerrückmeldungen zu erhöhen und somit günstigere Attributionsstile zu fördern.

Des Weiteren ist zu empfehlen, dass die Lehrpersonen nicht nur die Ergebnisse der Schülerrückmeldungen transparent machen, sondern auch die Maßnahmen, welche die Lehrperson aufgrund dieser ergriffen haben. Dies könnte dazu beitragen, dass die Schülerinnen und Schüler die Änderungen im Unterrichtshandeln besser wahrnehmen und sich keine enttäuschten Erwartungen einstellen.

Ferner sollte den Schülerinnen und Schülern die gestiegene Verantwortung für gelingenden Unterricht aufgrund der Teilnahme an einem Schüler-Kurzfeedback-Verfahren klar gemacht werden. Diese Maßnahmen könnten innerhalb eines Coachings noch gezielter eingesetzt werden.

Deshalb wäre von Interesse, wie ein begleitendes Coaching die Wirkung von Schüler-Kurzfeedback beeinflusst. Dies ist, unter Berücksichtigung der bestehenden Ausbildungsstrukturen, vor allem im Hinblick auf die praktische Lehrpersonenaus- und -weiterbildung interessant. Denn hier könnte das Schüler-Kurzfeedback-Verfahren im Zusammenspiel mit dem jeweiligen Mentor den Blick auf die Tiefenstrukturen des gehaltenen Unterrichts schärfen.

Ergänzend zu einem quantitativen Forschungsansatz wäre auch eine Vertiefung der qualitativen Forschung von Interesse. Denkbar sind beispielsweise kombinierte Schülerfeedback-Videostudien, die untersuchen, auf welche Weise sich das Lehrpersonenhandeln aufgrund des Schüler-Feedbacks aus der Perspektive von Experten ändert, um weitere Muster im Zusammenhang zwischen Schüler-Feedback, Unterrichtshandeln und deren Wirkmechanismen zu finden.

### **6.7.2 Folgerungen für die Schulpraxis**

Für die Schulpraxis bleibt generell festzuhalten, dass das Schüler-Kurzfeedback-Verfahren hilft, sich über die Ziele des Unterrichts bewusst zu werden und dass die Rückmeldungen der Schülerinnen und Schüler in den meisten Fällen ernst zu nehmen sind und häufig zu Veränderungen in der Unterrichtsentwicklung führen. Diese Veränderungen können sich in einer Verbesserung der Unterrichtsqualität widerspiegeln. Würden die bereits im vorigen Abschnitt dargestellten Muster während des Einsatzes des Verfahrens berücksichtigt, könnten sich die entsprechenden Wirkungen weiter erhöhen.

Aufgrund der Untersuchungsergebnisse kann der Einsatz dieses Schüler-Kurzfeedback-Verfahrens in der Praxis grundsätzlich empfohlen werden (bei Klassenverbänden mit schwerwiegenden sozialen Konflikten nur mit begleitenden Maßnahmen). Hierfür spricht auch, dass

fast alle befragten Lehrpersonen das Schüler-Feedback-Verfahren, wenn es in elektronischer Form vorläge, in ihrem Unterricht einsetzen würden, um das, in dieser Studie beschriebene Potential zur Unterrichtsentwicklung zu nutzen.

Die Entwicklung eines entsprechenden Programms wäre ein wichtiger Schritt, um das Lernverhalten der Schülerinnen und Schüler in arbeitsökonomischer Weise sichtbar zu machen und somit die Brücke aus der Wissenschaft in die Praxis zu schlagen, was wiederum der Wissenschaft neue Möglichkeiten eröffnen würde, sich mit dem Phänomen Schüler-Feedback vertiefter auseinanderzusetzen.

## 7 Literaturverzeichnis

- Abele, A. (1985). Thinking about Thinking. Causal, evaluative and finalistic cognitions about social situations. *European Journal of Social Psychology*, 15, S. 315-332.
- Abele, A. (1999). Motivationale Mediatoren von Emotionseinflüssen auf die Leistung: Ein vernachlässigtes Forschungsgebiet. In: M. Jerusalem & R. Pekrun (Hrsg.): *Emotion, Motivation und Leistung*. Goettingen, Bern, Toronto, Seattle: Hogrefe, S. 31-49.
- Adamina, M. & Möller, K. (2010). Zugänge zum naturwissenschaftlichen Lernen öffnen. In: P. Labudde (Hrsg.): *Fachdidaktik Naturwissenschaft*. Bern: Haupt Berne, S. 103-116.
- Ames, C. (1992). Classrooms: Goals, structures and student motivation. *Journal of Educational Psychology*, 84, S. 261-271
- Atwater, L. & Brett, R. (2005). Antecedents and consequences of reactions to developmental 360-degree feedback. *Journal of Vocational Behavior*. 66(3), S. 532-548.
- Bandura, A. (1986). Social foundations of thought and action. Englewood Cliffs: Prentice Hall.
- Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.
- Bandura, A. & Wood, R. (1989). Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology*, 56(5), S. 805-814.
- Bell, S. T. & Arthur, W. (2008). Feedback acceptance in developmental assessment centers: the role of feedback message, participant personality, and affective response to the feedback session. *Journal of Organizational Behavior*, 29(5), S 681-703.
- Bastian, J., Combe, A. & Langer R. (2001). Feedback-Methoden. Weinheim und Basel: Beltz.
- Berger, R. (2013). Warum fragt ihr nicht einfach uns?. Weinheim und Basel: Beltz.
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In: M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand et al. (Hrsg.): *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann, S. 314–354.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. In: *Educational Assessment, Evaluation and Accountability*, 21(1), S.5-13.
- Bless, H. & Fiedler, K. (1999). Förderliche und hinderliche Auswirkungen emotionaler Zustände auf kognitive Leistungen im sozialen Kontext. In: M. Jerusalem & R. Pekrun (Hrsg.): *Emotion, Motivation und Leistung*. Goettingen, Bern, Toronto, Seattle: Hogrefe, S. 9-29.
- Börlin, J. (2012). Das Experiment als Lerngelegenheit – Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität. In: H. Niedderer, H. Fischler, E. Sumfleth (Hrsg.): *Studien zum Physik- und Chemielernen*. Berlin: Logos Verlag.

- Börlin, J., Beerenwinkel, A. & Labudde, P. (2014). Bericht Analyse MINT-Nachwuchsbarometer-Auswertung der Datenerhebung vom Frühsommer 2012. Unter: [http://www.satw.ch/mint-nachwuchsbarometer/MINT-Nachwuchsbarometer\\_Bericht\\_Studie.pdf](http://www.satw.ch/mint-nachwuchsbarometer/MINT-Nachwuchsbarometer_Bericht_Studie.pdf) (abgerufen am 17.01.2017)
- Bohner, G., Bless, H., Schwarz, N. & Strack, F. (1988). What triggers causal attributions? The impact of valence and subjective probability. *European Journal of Social Psychology*, 18, S. 335-345.
- Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, S. 245-281.
- Cianci, A., Schaubroeck, J. & McGill, G. (2010). Achievement goals, feedback and task performance. *Human Performance*, 23(2), S. 131-154.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2. Auflage)*. Hillsdale, NJ: Erlbaum.
- Cohen, A., Doveh, E. & Eick, U. (2001). Statistical properties of the  $r_{WG(J)}$  index of agreement. *Psychological Methods*, 6, S. 297-310.
- Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A metaanalysis of findings. *Research in Higher Education*, 13, S. 321-341.
- Cornelius-White, J. (2007). Learner-Centered Teacher-Student Relationships Are Effective: A Meta-Analysis. *Review of Educational Research*, 77(1), S. 113-143.
- Cusella, L. P. (1982). The effects of source expertise and feedback valence on intrinsic motivation. *Human Communication Research*, 6(1), 17-32.
- Costa, P. & McCrea, R. (1992). *Revised NEO Personality Inventory (NEO PI-R) an NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa: Psychological Assessment Resources.
- Denzin, N. K. (1970). *The research act. A theoretical introduction to sociological methods*. Chicago: Aldine.
- Diehl, Joerg M. (2001). Studentische Lehrevaluation in den Sozialwissenschaften: Fragebögen, Normen, Probleme. In: E. Keiner (Hrsg.): *Evaluation (in) der Erziehungswissenschaft*. Weinheim, S. 63-90.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In: A. Helmke, W. Hornstein, E. Terhart, (Hrsg.): *Qualität und Qualitätssicherung im Bildungsbereich; Schule, Sozialpädagogik, Hochschule (Zeitschrift für Pädagogik, Beiheft; 41)*. Weinheim: Beltz, S. 73-92
- Ditton, H. (2002). Lehrkräfte und Unterricht aus Schülersicht. Ergebnisse einer Untersuchung im Fach Mathematik. *Zeitschrift für Pädagogik*, 48(2), S. 262-286.

- Ditton, H. & Arnoldt, B. (2004). Wirksamkeit von Schülerfeedback zum Fachunterricht. In: J. Doll, & M. Prenzel (Hrsg.): *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung*. Münster: Waxmann, S. 152-170.
- Ditton, H., Arnoldt, B., Bornemann, E. (2002). Entwicklung und Implementation eines extern unterstützten Systems der Qualitätssicherung an Schulen – QuaSSU. M. Prenzel & J. Doll (Hrsg.): *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen*. Weinheim: Beltz 2002, S. 374-389.
- Ditton, H. & Müller, A. (Hrsg.) (2014): *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder*. Münster: Waxmann.
- Donovan, J. & Hafsteinsson, L. (2006). The impact of goal-performance discrepancies, self efficacy and goal orientation on upward goal revision. *Journal of Applied Social Psychology*, 36(4), S. 1046-1069.
- Duit, R. (1996). Lernen als Konzeptwechsel im naturwissenschaftlichen Unterricht. In: R. Duit & C. v. Rhöneck (Hrsg.): *Lernen in den Naturwissenschaften. Beiträge zu einem Workshop an der Pädagogischen Hochschule Ludwigsburg*. Kiel, S. 145-162.
- Duit, R., Tesch, M. & Mikelskis-Seifert, S. (2010). Das Experiment im Physikunterricht. Piko-Brief Nr. 7. Unter: <http://www.ipn.uni-kiel.de/de/das-ipn/abteilungen/didaktik-der-physik/piko/pikobriefe032010.pdf> (abgerufen am 13.10.2016)
- Dweck, C. (1989). Motivation. In: A. Lesgold & R. Glaser (Hrsg.): *Foundations of a psychology education*. Hillsdale: Erlbaum, S. 87-135.
- EDK (2011). *Grundkompetenzen für die Naturwissenschaften. Nationale Bildungsstandards*. Schweizerische Konferenz der kantonalen Erziehungsdirektoren. Unter: [http://edudoc.ch/record/96787/files/grundkomp\\_nawi\\_d.pdf](http://edudoc.ch/record/96787/files/grundkomp_nawi_d.pdf) (abgerufen am 20.01.2017)
- Eikenbusch, E. (2001). Erfahrungen mit Schülerrückmeldung in der Oberstufe. *Pädagogik*, 53(5), S.18-22.
- Elliot, E. & Dweck, C. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), S. 5-12.
- Feys, M., Anseel, F. & Wille, B. (2011). Improving feedback reports: The role of procedural information and information specificity. *Academy of Management Learning & Education*, 10(4), S. 661-681.
- Fielding N. G., Fielding J. L. (1986). Linking Data. *Qualitative Research Methods* (4). Thousand Oaks, London, New Delhi: SAGE Publications.
- Fischler, H. (1989). Orientations of the Actions of Science Teachers. *International Journal of Science Education* 11 (2), S. 185-193.

- Flick, U. (2000). Triangulation in der qualitativen Forschung. In: U. Flick, E. von Kardoff & I. Steinke (Hrsg.): *Qualitative Forschung. Ein*. Reinbek bei Hamburg: Rowohlt. S. 309-316.
- Flick, U. (2011). *Triangulation – Eine Einführung*. VS Verlag Wiesbaden.
- Gärtner, H. (2013). Wirksamkeit von Schülerfeedback als Instrument der Selbstevaluation von Unterricht. In: J. Hense, S. Rädiker, W. Böttcher, T. Widmer (Hrsg.), *Forschung über Evaluation: Bedingungen, Prozesse, Wirkungen*. Münster: Waxmann, S. 107-124.
- Gärtner, H. & Vogt, A. (2013). Wie Lehrkräfte Ergebnisse eines Schülerfeedbacks verarbeiten und nutzen. *Unterrichtswissenschaft*, 3, S. 252-267.
- Geller, C. Neumann, K., & Fischer, H. E. (2014). A deeper Look inside Teaching Scripts: Learning Process Orientations in Finland, Germany and Switzerland. In: H. E. Fischer, P. Labudde, K. Neumann, Jouni Viiri (Hrsg.): *Quality of Instruction in Physics. Comparing Finland, Germany and Switzerland*. Münster, New York: Waxmann, S. 81-92.
- Gediga, G., Kannen, K. v., Schneider, F., Köhne, S., Luck, H., & Schneider, B. (2000): *KIEL: Ein Kommunikationsinstrument für die Evaluation von Lehrveranstaltungen*, unter: <http://www.methodos.de/books/kiel.pdf> (abgerufen am 06.10.2016).
- Gerstenmaier, J. (1975). *Urteile von Schülern über Lehrer: Eine Analyse ausgewählter empirischer Untersuchungen*. Weinheim: Beltz.
- Giffin, , K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin*, 68(2), S. 104-120.
- Gläser-Zikuda, M., Seidel, T., Rohlf, C., Gröschner A. & Ziegelbauer, S. (2012) (Hrsg.). *Mixed Methods in der empirischen Bildungsforschung – eine Einführung in die Thematik*. In: *Mixed Methods in der empirischen Bildungsforschung*. Münster: Waxmann. S.7-13.
- Gläser-Zikuda, M. & Järvelä, S. (2008) (Hrsg.). Qualitative and quantitative approaches to enrich understanding of emotional and motivational aspects of learning. *International Journal of Educational Research*, 47, S.79-83.
- Glaserfeld, E. v. (1992). Konstruktion der Wirklichkeit und des Begriffs der Objektivität. In: H. Gumin & H. Meier (Hrsg.): *Einführung in den Konstruktivismus*. München und Zürich: Piper. S. 9-39.
- Gräsel, C. & Göbel, K. (2011). Unterrichtsqualität. In: H. Reinders, H. Ditton, C. Gräsel, & B. Gniewosz: *Empirische Bildungsforschung*- Wiesbaden: Springer VS. S. 87-98.
- Graf, B (2001). Verbesserungen des Unterrichts gemeinsam planen und umsetzen. *Pädagogik*, 53(5), S. 10-13.
- Grob, R. (in Druck). *Towards the implementation of formal formative assessment in inquiry-based science education in Switzerland*.

- Gröschner, A. & Kleinknecht, M. (2013). Qualität von Unterricht – Ansätze aus der Perspektive der Unterrichtsforschung. In: L. Haag, S. Rahm, H. J. Apel & W. Sacher (Hrsg.): *Studienbuch Schulpädagogik*. Bad Heilbrunn: Verlag Julius Klinkhardt. S. 162- 177.
- Gröschner, A., Seidel, T. & Shavelson, R.J. (2013). Methods for Studying Teacher and Teaching Effectiveness. In: J. Hattie & E.M. Anderman (Hrsg.): *International Guide of Student Achievement*. London/New York: Routledge. S. 240-243.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen - Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Haladyna, T. & Hess, R. K. (1994). The detection and correction of bias in student ratings of instruction. *Research in Higher Education*, 35, S. 669-687.
- Halperin, K., Snyder, C. R., Shenkel, R. J. & Houston, B. K. (1976). Effects of source status and message favorability on acceptance of personality feedback. *Journal of Applied Psychology*, 61(1), S. 85-88.
- Hascher, T. (2005). Emotionen im Schulalltag: Wirkungen und Regulationsformen. *Zeitschrift für Pädagogik* 51(5), S. 610-625.
- Hattie, J. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), S. 81-112
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer/Klett.
- Helmke, A. & Hosenfeld, I. (2005). Standardbezogene Unterrichtsevaluation. In G. B. Brägger, B. & Landwehr, N. (Hrsg.), *Schlüsselfragen zur externen Schulevaluation*. Bern: hep Verlag, S. 127-151.
- Helmke, A., Schneider, W. & Weinert, F. E. (1986). Quality of instruction and classroom learning outcomes: The German contribution to the IEA classroom environment study. *Teaching and Teacher education*, 2, S. 1-18.
- Hempel, P. (2008). Chinese Reactions to performance feedback: Non-task attributions of feedback intentions. *Asia Pacific Journal of Human Resources*, 46(2), S. 196-219.
- Helmke, A. & Lenske, G. (2013). Unterrichtsdiagnostik als Voraussetzung für Unterrichtsentwicklung. *Beiträge zur Lehrerbildung*, 31(2), S. 214-233.
- Helmke, A., Helmke, T., Lenske, L., Pham, G. H., Praetorius, A.-K., Schrader, F.-W. & Ade-Thurow, M. (2014). Unterrichtsdiagnostik mit EMU (Evidenzbasierte Methoden der Unterrichtsdiagnostik und -entwicklung). In: M. Ade-Thurow, W. Bos, A. Helmke, T. Helmke et al. (Hrsg.): *Aus- und Fortbildung der Lehrkräfte in Hinblick auf Verbesserung der Diagnosefähigkeit, Umgang mit Heterogenität und individuelle Förderung*. Münster: Waxmann. S. 150-163.



- Herrmann, J., Höfer, C. (1999). Evaluation in der Schule - Unterrichtsevaluation: Berichte und Materialien aus der Praxis. Gütersloh: Verlag Bertelsmann Stiftung.
- Heslin, P. & Latham, G. (2004). The effect of an upward feedback on managerial behavior. *Applied Psychology*, 53(1), S. 23-37.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), S. 99-107.
- Hochschulrektorenkonferenz (2010). *Beiträge zur Hochschulpolitik 8/2010: Wegweiser 2010 – Qualitätssicherung an Hochschulen; Projekt Qualitätsmanagement*, Bonn.
- Hosenfeld, I., Groß Ophoff, J. (2007). Nutzung und Nutzen von Evaluationsstudien in Schule und Unterricht. *Empirische Pädagogik*, 21(4), S. 352-367.
- Hofer, M. (1981): Lehrerverhalten aus der Sicht der Schüler. *Pädagogische Welt* 35, S. 49-56.
- Hook, C. M. & Rosenshine, B. V. (1979). Accuracy of teacher reports of their classroom behavior. *Review of Educational Research*, 49, S. 1-12.
- Ilgen, D. & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Journal of applied Psychology*, 48(3), S. 550-565.
- Ilgen, D., Fisher, C. & Taylor, S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), S. 349-371.
- Isen, A. M. ; Daubman, K. A. & Nowicki & Gary P. (1987). Positive Affect Facilitates Creative Problem Solving. *Journal of Personality and Social Psychology*, 52(6), S. 1122-1131.
- Isen, A. M. & Means, B. (1983). The influence of positive affect on decision-making strategy. *Social Cognition*. 2, S. 18-31.
- John, O. P. & Robins, R. W. (1993). Determinants of interjudge agreement of personality traits: The big five domains, observability, evaluativeness and the unique perspective of the self. *Journal of Personality*, 61, S. 521-551.
- Junge, C., Arx, M. von, Labudde, P. (2014). Classroom Management. In: H. E. Fischer, P. Labudde, K. Neumann, Jouni Viiri (Hrsg.): *Quality of Instruction in Physics. Comparing Finland, Germany and Switzerland*. Münster, New York: Waxmann, S. 161-176.
- Kenny, D. (2004) PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review*, 8 (3), S. 265-280.
- Kircher, E., Girdwidz, R., Häußler, P. (2007). *Physikdidaktik*. Berlin, Heidelberg: Springer-Verlag.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), S. 75-86.

- Klieme, E., Schümer, G. & Knoll, S. (2001): Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung im internationalen Vergleich. In: E. Klieme, & J. Baumert (Hrsg.): *TIMSS – Impulse für Schule und Unterricht*. Bonn: BMBF. S. 43-57.
- Kluger, A & Denisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), S. 82-94.
- KMK-Bildungsstandards (2004). Unter: <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsstandards.html> (abgerufen am: 18.10.2016)
- Koch, A. (in Druck). *Naturwissenschaftliches Lehren neu gestalten: Analysen zu Determinanten und Verläufen handlungswirksamer Unterrichtsentwicklung in der obligatorischen Schule*.
- Koziol, S. M. & Burns, P. (1986). Teachers' accuracy in self-reporting about instructional practices using a focused self-report inventory. *Journal of Educational Research*, 79 (4), S. 205-209.
- Krasmann, J. (2011). The feedback-seeking personality: Big five and feedback-seeking behavior. *Journal of Leadership & Organizational Studies*, 17(1), S. 18-32.
- Kretschmer, W (2013): Schülerfeedback und Hospitationsringe als verbundene Werkzeuge – ein Praxisbeispiel. In: C.G. Buhren (Hrsg.): *Handbuch Feedback in der Schule*. Weinheim und Basel: Beltz.
- Kuckartz, U. (2014). *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*. Wiesbaden: Springer.
- Kuhn, J., Müller, A., Müller, W., & Vogt, P. (2010). Kontextorientierter Physikunterricht. In: *Praxis der Naturwissenschaften - Physik in der Schule*, S. 13-25.
- Kulhavy, R. W. & Stock, W. A. (1989). Feedback in written instruction: The place of response certainty. *Educational Psychology Review*, 1, S. 279-308.
- Kulhavy R. W. & Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. In: J. V. Dempsey & G. C. Sales (Hrsg.): *Interactive instruction and feedback*. Englewood Cliffs, NJ: Educational Technology Publications, S. 3-20.
- Kulhavy R. W., White, M. T., Topp, B. W., Chan, A. L. & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, 10, S. 285-291.
- Kunter, M & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In: M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.): *Professionelle Kompetenz von Lehrkräften Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann. S. 85-113.
- Lance, C. E., LaPointe, J. A. & Stewart, H. M. (1994). A test of the context dependency of three causal models of halo rater error. *Journal of Applied Psychology*, 79, S. 332-340.
- Lang, J. W. B. & Kersting, M. (2007). Langfristige Effekte von regelmäßigem Feedback aus

- studentischen Lehrveranstaltungsevaluationen. In: A. Kluge & K. Schüler (Hrsg.): *Qualitätssicherung und -entwicklung an Hochschulen: Methoden und Ergebnisse*. Lengerich: Pabst Science Publishers, S. 159-167.
- Landwehr, N. (2003/2007). *Grundlagen zum Aufbau einer Feedbackkultur*. Bern: h.e.p. Verlag.
- Lenske, G. (2016). *Schülerfeedback in der Grundschule*. Münster: Waxmann.
- Leung, K., Su, S. & Morris, M. (2001). When is criticism not constructive? The roles of fairness perceptions and dispositional attributions in employee acceptance of critical supervisory feedback. *Human Relations*, 54(9), S. 1155-1187.
- L'Hommedieu, R., Menges, R. J. & Brinko, K. T. (1990). Methodological explanations for the modest effect of feedback from student ratings. *Journal of Educational Psychology*, 82, S. 232-241.
- Lipowsky, F. (2002). Zur Qualität offener Lernsituationen im Spiegel empirischer Forschung – Auf die Mikroebene kommt es an. In: U. Drews & U. Wallrabenstein (Hrsg.): *Freiarbeit in der Grundschule. Offener Unterricht in Theorie, Forschung und Praxis*. Frankfurt a. M.: Grundschulverband. S. 126-159.
- Lotz, M. (2016). *Kognitive Aktivierung im Leseunterricht der Grundschule – Eine Videostudie zur Gestaltung und Qualität von Leseübungen im ersten Schuljahr*. Wiesbaden: Springer.
- Lyden, J., Chaney, L., Danehower, V. & Houston, D. (2002). Anchoring, attributions and self-efficacy: An examination of interactions. *Contemporary Educational Psychology*, 27(1), S. 99-117.
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten. Ansätze zur Bestimmung der Reliabilität und Übereinstimmung von Schülerwahrnehmungen. *Zeitschrift für Pädagogische Psychologie*, 20 (1/2). S. 85-96.
- Lüdtke, O., Trautwein, U., Schnyder, I. & Niggli, A. (2007). Simultane Analysen auf Klassenebene. Eine Demonstration der konfirmatorischen Mehrebenen-Faktorenanalyse zur Analyse von Schülerwahrnehmungen am Beispiel der Hausaufgabenvergabe. *Zeitschrift für Pädagogische Psychologie*, 39 (1). S. 1-11.
- Marsh, H. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Hrsg.): *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Netherlands: Springer, S. 319-383.
- Mayring, P. (2001). Kombination und Integration qualitativer und quantitativer Analyse. *Forum Qualitative Sozialforschung (Online Journal)* 2 (1). Unter: <http://www.qualitative-research.net/index.php/fqs/article/view/967/2111> (abgerufen am: 10.11.2016).
- Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlage und Techniken*. Weinheim: Beltz.

- McCrea, R. & Costa, T. (2011). The five-factor-theory of personality. In: O. John, R. Robins & L. Pervon (Hrsg.): *Handbook of personality: Theory and research*. New York: Guilford, S. 159-181.
- Merzyn, G. (2008). *Naturwissenschaften, Mathematik, Technik – immer unbeliebter?* Baltmannsweiler: Schneider Verlag Hohengehren.
- Möller, K. (2007). Genetisches Lernen und Conceptual Change. In: J. Kahlert et al. (Hrsg.): *Handbuch Didaktik des Sachunterrichts*. Bad Heilbrunn: Klinkhardt, S. 258-266.
- Möller, K. (2010). Lernen von Naturwissenschaften heisst: Konzepte verändern. In: P. Labudde (Hrsg.): *Fachdidaktik Naturwissenschaft*. Bern: Haupt Berne, S. 57-53.
- Mory, E. H. (2004). Feedback research revisited. In: D. H. Jonassen (Hrsg.): *Handbook of research on educational communications and technology* (2. Ausgabe). Mahwah, NJ: Lawrence Erlbaum Associates, S. 745-783.
- Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. Reihe Pädagogische Psychologie und Entwicklungspsychologie, Band 56. Münster: Waxmann.
- Narciss, S. (2014). Modelle zu den Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen. In: H. Ditton & A. Müller (Hrsg.): *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder*. Münster: Waxmann, S. 43-82.
- Oberhoff, B. (1978). *Akzeptanz von interpersonellem Feedback*. Münster: Philosophische Fakultät der Universität Münster.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3. Aufl.). Thousand Oaks: Sage.
- Pawek, C. (2009). *Schülerlabore als interessefördernde auserschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe*. Unter: [http://www.dlr.de/schoollab/Portaldata/24/Resources/dokumente/Diss\\_Pawek.pdf](http://www.dlr.de/schoollab/Portaldata/24/Resources/dokumente/Diss_Pawek.pdf) (abgerufen am: 10.02.2016)
- Phye, G. D. & Bedner, T. (1989). Feedback complexity and practice: Response pattern analysis in retention and transfer. *Contemporary Educational Psychology*, 14, S. 97-110.
- Piaget, J. (1970). *Science of Education and the Psychology of the Child*. New York: Orion Press.
- Reusser, K. (2006). Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In M. Baer, M. Fuchs, P. Füglistner, K. Reusser & H. Wyss (Hrsg.), *Didaktik auf psychologischer Grundlage: Von Hans Aeblis kognitionspsychologischer Didaktik zur modernen Lehr- und Lernforschung*. Bern: h.e.p, S. 151- 168.
- Rindermann, H. (2009). *Lehrevaluation. Einführung und Überblick zu Forschung und Praxis der*

- Lehrveranstaltungsevaluation an Hochschulen mit einem Beitrag zur Evaluation computerbasierten Unterrichts* (2., leicht korrigierte Aufl.). Landau: Empirische Pädagogik.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In: L. Berkowitz (Hrsg.): *Advances in experimental social psychology* (Volume 10). New York: Academic Press. S. 173-220.
- Sadler (1989) siehe visible learning
- Schmidt, B. (2008). Warum oft wirksam? Und warum manchmal wirkungslos? – Subjektive Erklärungen zur Wirkung von Lehrveranstaltungsevaluation aus Sicht von Nutzern und Anbietern. *Zeitschrift für Evaluation*, 7(1), S. 7-33.
- Seidel, T., Prenzel, M., Dalhefte, I. M., Meyer, L. Trepke, C., Lehrke, M., & Duit, R. (2003). Überblick über die IPN-Videostudie in Physik. In: T. Seidel, M. Prenzel, R. Duit & M. Lehrke (Hrsg.): *Technischer Bericht zur Videostudie "Lehr-Lernprozesse im Physikunterricht"*. Kiel: IPN.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), S. 454-499.
- Sinclair, R. C. (1988). Mood, categorization breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations. *Organizational Behavior and Human Decision Process*, 42, S. 22-46.
- Slavin, Robert E. (1996). *Education for all*. Lisse: Swets & Zeitlinger.
- Semmer, N. & Jacobshagen, N. (2010). Feedback im Arbeitsleben – eine Selbstwert-Perspektive. *Gruppendynamik Organisationberatung*, 41(1), S. 39-55.
- Smither, J. W., London, M. & Richmond, R. R. (2005a). The relationship between leaders personality and their reactions to and use of multisource feedback: A longitudinal study. *Group and Organization Management*, 30(2), S. 181-210.
- Stajkovic, A. & Sommer, S. (2000). Self efficacy and work-related performance: A meta analysis. *Psychological Bulletin*, 124(2), S. 240-261.
- Strijbos, J. W. & Müller, A. (2014). Personale Faktoren im Feedbackprozess. In: H. Ditton & A. Müller (Hrsg.): *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder*. Münster: Waxmann. S. 83-134.
- Steelman, L. A., Levy, P. E. & Snell, A. F. (2004). The feedback environment scale: Construct definition, measurement, and validation. *Educational and Psychological Measurement*, 64(1), S. 165-184.
- Strijbos, J. W., Narciss, S. & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: are they critical für feedback perceptions and efficiency? *Learning and instruction*, 30, S. 291-303.

- Tashakkori, A. & Teddlie, C. (Hrsg.) (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks: Sage.
- Tesch, M. & Duit, R. (2004). Experimentieren im Physikunterricht - Ergebnisse einer Videostudie. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, S. 51-59.
- Tippelt, R. & Schmidt, H. (2010). *Handbuch Bildungsforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Veenman, S., Kenter, B., & Post, K. (2000). Cooperative learning in Dutch primary classrooms. *Educational Studies*, 26(3), S. 281-302.
- Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Wagner, W. (2008). *Methodenprobleme bei der Analyse der Unterrichtswahrnehmung aus Schülersicht – am Beispiel der Studie DESI (Deutsch Englisch Schülerleistungen International) der Kultusministerkonferenz*. Unter: <http://d-nb.info/987591800/34> (abgerufen am 12.04.2016).
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), S. 249-294.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. In: *Psychological Review*, 92(4), S. 548-573.
- Wiesner, H. (2007). Verbesserung des Lernerfolgs im Unterricht über Mechanik. In: R. Müller, R. Wondzinski, & Hopf, Martin (Hrsg.): *Schülervorstellungen in der Physik*. Bad Honnef: Aulis Verlag Deubner, S. 114-119.
- Wilde, M., Bätz, K., Kovaleva A. & Urhahne, D. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, S. 31-45.
- Wodzinsky, R. (2007). Lernschwierigkeiten in der Mechanik. In R. Müller, R. Wodzinski, & M. Hopf, *Schülervorstellungen in der Physik*. Köln: Aulis Verlag Deubner, S. 114-119.
- Wood, D Bruner, J. S. & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, 17(2), S. 89-100.
- Zöller, A (2016). Modellprojekt Schüler-Feedback in der 2. Phase der Lehrerbildung. In: *Staatsinstitut für Schulqualität und Bildungsforschung*. Unter: [https://www.isb.bayern.de/download/19038/16\\_11\\_07\\_leitfaden\\_schuelerfeedback.pdf](https://www.isb.bayern.de/download/19038/16_11_07_leitfaden_schuelerfeedback.pdf)

## 8 Anhang

### 8.1 Erhebungsinstrumente

#### 8.1.1 Der Schüler-Kurzfeedback-Bogen

##### Liebe/r xxx

in diesem Fragebogen findest Du Fragen über:

- Deine Wahrnehmungen, Dein Denken und Verhalten im heutigen Physikunterricht

Bitte lies jede Frage sorgfältig durch und beantworte sie so genau wie möglich.

In diesem Fragebogen musst Du als Antwort meistens ein Kästchen ankreuzen. Hinzu kommen zwei Fragen, die du mit einem kurzen Text beantworten kannst. Wenn Du beim Ankreuzen einen Fehler machst, male das angekreuzte Kästchen ganz aus und kreuze die richtige Antwort an. Wenn Du etwas Falsches aufgeschrieben hast, streiche es einfach durch und schreibe die neue Antwort daneben.

**In diesem Fragebogen gibt es keine richtigen oder falschen Antworten. Gib die Antworten an, die für Dich am besten passen.**

Wenn Du etwas nicht verstehst oder nicht weißt, wie Du antworten sollst, darfst Du deine/n Lehrer/in um Hilfe bitten.

##### Fragen zum heutigen Unterricht (pro Zeile nur 1 Kreuz)

	Trifft voll zu	Trifft größten- teils zu	Trifft teilweise zu	Trifft gar nicht zu
Ich bin heute im Unterricht nicht mitgekommen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich finde, dass man den Unterrichtsstoff schneller hätte durchführen können.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Unterricht hat mich dazu angeregt, genauer über das Unterrichtsthema nachzu- denken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Für mich war klar, welche Inhalte im heutigen Unterricht gelernt werden sollen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühlte mich heute von den Mitschülern/innen gestört.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand den Unterrichtsinhalt zu schwer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand den Unterrichtsinhalt zu leicht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Unterricht hat dazu geführt, dass ich mein bisheriges Wissen über das heutige Thema korrigiert oder erweitert habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Für mich war klar, welches Lernziel heute erreicht werden sollte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Um wirklich gut lernen zu können, war es heute im Unterricht zu laut.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich hätte heute mehr Unterstützung gebraucht, um im Unterricht mitzukommen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Heute wurden Dinge, die eigentlich klar waren, zu ausführlich besprochen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe heute in Physik etwas dazugelernt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin heute Schritt für Schritt dem Lernziel näher gekommen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe mir heute gewünscht, dass die Lehrperson früher für Ordnung gesorgt hätte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mir fiel es heute schwer, dem Unterricht zu folgen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand, dass es heute im Unterricht zu langsam vorwärts ging.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe mich heute intensiv mit den Unterrichtsinhalten beschäftigt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Für mich ist klar, wie die Inhalte des heutigen Unterrichtsinhalts zusammenhängen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Heute war im Unterricht eine gute Atmosphäre.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Reihenfolge mit der die heutigen Inhalte durchgenommen wurden, verwirrte mich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Lehrperson hatte die Klasse heute gut im Griff.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Alle Deine Antworten werden vertraulich behandelt.**

### Offene Fragen

Folgende Themen des heutigen Unterrichts sollten nochmals angesprochen werden, damit ich sie verstehe:

---



---

Was ich unbedingt noch zu der heutigen Schulstunde sagen will:

---



---

Falls heute Experimente und/oder einfache Versuche durchgeführt wurden, beantwortet bitte die folgenden Fragen.

Auch wenn in den folgenden Fragen von Experimenten (Mehrzahl) die Rede ist, sollt ihr die Fragen auch beantworten, wenn nur ein Experiment durchgeführt wurde. Es sind hier Experimente gemeint, die von euch selbst oder vom Lehrer durchgeführt wurden. Auch ganz einfache und kurze Versuche gelten als Experiment. Wenn mehrere Experimente gemacht wurden, dann gebe bei den Antworten einen "Durchschnittswert" an.

Wie viele Experimente oder Versuche wurden heute insgesamt gemacht? \_\_\_\_\_

Fragen zu Experimenten (pro Zeile 1 Kreuz)	Trifft voll zu	Trifft größten- teils zu	Trifft teilweise zu	Trifft gar nicht zu	Kann ich nicht beantwor- ten
Die heutigen Experimente haben mich dazu angeregt über den Unterrichtsinhalt genauer nachzudenken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ich habe während des Experimentierens darauf geachtet, ob genau gearbeitet wird.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Die heutigen Experimente haben mir geholfen, die Herleitung der Theorie oder die Theorie selber, besser zu verstehen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ich habe darauf geachtet, ob die Regeln zum richtigen Experimentieren eingehalten werden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Experimente haben dazu geführt, dass ich mein bisheriges Wissen über das heutige Thema korrigiert oder erweitert habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Während der Experimente habe ich mir überlegt, wie man vorgehen muss, um eine Hypothese (Vermutung) eindeutig zu überprüfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ich habe mir Gedanken gemacht, wie die Ergebnisse der Experimente zu erklären sind.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	



## 8.1.2 Schüler-Fragebogen Präerhebung

		Stimme voll zu	stimme eher zu	Stimme eher nicht zu	stimme gar nicht zu
		4	3	2	1
<b>1.</b>	<b>Wie beurteilst du folgende Aussagen zur Verständlichkeit des Unterrichts?</b> (Pro Zeile bitte nur ein Kreuz.)				
	Unsere Lehrerin/unser Lehrer...				
V1SKLA02	drückt sich klar und verständlich aus.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SKLA01	spricht laut und deutlich, so dass man ihn gut versteht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SKLA03	stellt wichtige Sachverhalte deutlich heraus.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>2.</b>	<b>Wie beurteilst du folgende Aussagen zum Unterricht?</b> (Pro Zeile bitte nur ein Kreuz.)				
	Unsere Lehrerin/unser Lehrer...				
V1SQIS03	gibt vorab einen Überblick zur Gliederung des Stoffes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQFS02	geht im Unterricht in einer logischen Reihenfolge vor.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQFS03	macht Übergänge zu neuen Themenbereichen deutlich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQFS04	stellt die Beziehungen zwischen behandelten Themen heraus.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQIS04	fasst abschließend die wichtigsten Inhalte und Ergebnisse zusammen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQIS02	verwendet Übersichten, um Zusammenhänge aufzuzeigen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQFS01	verweist auf Zusammenhänge mit schon durchgenommenem Stoff.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQIS01	stellt Zusammenhänge mit dem Stoff anderer Fächer her.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SQFS05	zeigt bei Aufgaben die Vor- und Nachteile verschiedener Lösungswege auf.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>3.</b>	<b>Wie weit stimmst du den Aussagen zum Unterrichtstempo zu?</b> (Pro Zeile bitte nur ein Kreuz.)				
v1STEM03	Wir kommen mit dem Unterrichtsstoff nur langsam voran.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
v1STEM01	Es wird so viel Stoff wiederholt, dass wir kaum weiterkommen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
v1STEM02	Unsere Lehrerin/unser Lehrer geht so langsam im Stoff voran, dass es langweilig ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>4.</b>	<b>Wie beurteilst du die folgenden Aussagen zur Schwierigkeit des Unterrichtsstoffs?</b> (Pro Zeile bitte nur ein Kreuz.)				
V1SSCH05	Die Hausaufgaben sind so leicht, dass ich gar nicht wirklich nachdenken muss.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SSCH04	Unsere Lehrerin/unser Lehrer stellt sehr hohe Anforderungen im Unterricht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SSCH03	Der Unterricht ist so schwer, dass ich nicht mitkomme.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SSCH02	Die Aufgaben, die uns die Lehrerin/der Lehrer stellt, sind ganz schön schwierig.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SSCH01	Die Hausaufgaben sind oft so schwer, dass ich sie nicht lösen kann.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		Stimme voll zu	stimme eher zu	Stimme eher nicht zu	stimme gar nicht zu
		4	3	2	1
<b>5.</b>	<b>Merkt eure Lehrerin/euer Lehrer, wie ihr mit ihrem/seinem Unterricht zu Recht kommt? (Pro Zeile bitte nur ein Kreuz.)</b>				
	Unsere Lehrerin/unser Lehrer...				
V1SADK06	kann sich gut auf unsere Klasse einstellen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK11	geht weiter nach seinem Unterrichtsplan vor, auch wenn keiner mehr mitkommt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK10	merkt, wenn der Unterricht zu schwer ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK09	merkt, wenn wir uns im Unterricht langweilen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK08	merkt, wenn es Probleme in der Klasse gibt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK12	kennt die Stärken und Schwächen der einzelnen Schülerinnen und Schüler.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK04	weiß, wie er mit jedem von uns umzugehen hat.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK03	merkt sofort, wenn jemand nicht mehr mitkommt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>6.</b>	<b>Wie weit geht eure Lehrerin/euer Lehrer auf einzelne Schülerinnen und Schüler ein? (Pro Zeile bitte nur ein Kreuz.)</b>				
	Unsere Lehrerin/unser Lehrer...				
V1SVER06	kümmert sich um jede/n einzelne/n Schüler/in.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK07	gibt einzelnen Schülerinnen und Schülern Tipps, wie sie besser lernen können.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK01	nimmt sich Zeit, um einzelnen Schülerinnen und Schülern Dinge, die sie nicht verstanden haben, nochmals zu erklären.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK02	bemüht sich, dass alle im Unterricht mitkommen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SADK05	erklärt den Stoff so oft, bis es alle verstanden haben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>7.</b>	<b>In welchem Maße bemüht sich eure Lehrerin/euer Lehrer, den Unterricht interessant zu machen? (Pro Zeile bitte nur ein Kreuz.)</b>				
	Unsere Lehrerin/unser Lehrer...				
V1SINT05	verwendet Beispiele, um uns den Stoff klarzumachen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SINT07	gestaltet den Unterricht immer nach demselben Schema.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SINT06	gestaltet den Unterricht abwechslungsreich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SINT04	stellt uns interessante Aufgaben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SINT03	gestaltet den Unterricht zeitweise richtig spannend.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SINT01	benutzt Abbildungen, um den Lehrstoff zu veranschaulichen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SINT02	stellt Verbindungen zwischen dem Unterrichtsstoff und dem täglichen Leben her.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		Stimme voll zu	Stimme eher zu	Stimme eher nicht zu	Stimme gar nicht zu
		4	3	2	1
<b>8.</b>	<b>Wie weit stimmst du folgenden Aussagen zum Unterricht zu? (Pro Zelle bitte nur ein Kreuz.)</b>				
V18INT08	Unsere Lehrerin/unser Lehrer erklärt uns, warum das wichtig ist, was wir an Stoff behandeln.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18INT09	Unsere Lehrerin/unser Lehrer zeigt auf, was wir später mit dem, was wir im Unterricht lernen, anfangen können.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>9.</b>	<b>Wie beurteilst du das Verhältnis zu eurem Lehrer in eurer Klasse? (Pro Zelle bitte nur ein Kreuz.)</b>				
V18VER04	Wir haben ein vertrauensvolles Verhältnis zu unserer Lehrerin/unserem Lehrer.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18VER02	Unsere Lehrerin/unser Lehrer nimmt unsere Probleme ernst.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18VER05	Unsere Lehrerin/unser Lehrer bevorzugt im Unterricht einige Schülerinnen und Schüler.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18VER03	Unsere Lehrerin/unser Lehrer bemüht sich, auf unsere Wünsche einzugehen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18VER01	Wenn uns etwas nicht gefällt, können wir offen mit unserer Lehrerin/ unserem Lehrer darüber reden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>10.</b>	<b>Wie gut kann sich eure Lehrerin/euer Lehrer im Unterricht durchsetzen? (Pro Zelle bitte nur ein Kreuz.)</b>				
	Unsere Lehrerin unser Lehrer...				
V18MAN01	hat die Klasse im Griff.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18MAN02	schaft es nicht, im Unterricht für Ruhe und Ordnung zu sorgen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18MAN03	kann sich in unserer Klasse nicht durchsetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18MAN04	ist leicht vom Stoff abzulenken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		sehr oft	oft	sehr selten	nie
<b>11.</b>	<b>Wie häufig kommt es vor, dass ... (Pro Zelle bitte nur ein Kreuz)</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
V18ZEI01	eure Lehrerin/euer Lehrer zu spät zum Unterricht kommt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI02	Schülerinnen und Schüler zu spät zum Unterricht kommen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI03	die Lehrerin/der Lehrer während des Unterrichts weg muss?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI04	der Unterricht ausfällt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI05	ein großer Teil der Stunde schon vorbei ist, bevor ihr mit dem Unterricht richtig anfängt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI06	ihr schon lange vor der Pause mit dem Unterricht aufhört?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI07	die Lehrerin/der Lehrer Geschichten erzählt, die mit dem Stoff nichts zu tun haben?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18ZEI08	es im Unterricht Leerlaufzeiten gibt, in denen nichts getan wird?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18MAN05	es im Unterricht drunter und drüber geht?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		Stimme voll zu	Stimme eher zu	Stimme eher nicht zu	Stimme gar nicht zu
		4	3	2	1
<b>12.</b>	<b>Inwiefern stimmst du folgenden Aussagen über deine Lehrerin/deinen Lehrer zu? (Pro Zelle bitte nur ein Kreuz.)</b>				
V19PO801	Insgesamt ist unsere Lehrerin/unser Lehrer eine wirklich gute Lehrkraft.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19PO802	Insgesamt macht unsere Lehrerin/unser Lehrer einen wirklich guten Unterricht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19PO804	Bei dieser Lehrerin/diesem Lehrer fällt es mir leicht zu lernen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19PO805	Bei dieser Lehrerin/diesem Lehrer gehe ich gern in den Unterricht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19PO806	Diese Lehrerin/diesen Lehrer hätte ich auch gerne im nächsten Schuljahr.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19PO803	Wenn unser Lehrer/ unsere Lehrerin Schwierigkeiten bekäme, würde ich mich für ihn/sie einsetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>13.</b>	<b>Wie geht eure Lehrerin/euer Lehrer mit Antworten auf eure Fragen um? (Pro Zelle bitte nur ein Kreuz.)</b>				
	Unsere Lehrerin unser Lehrer...				
V19SHL01	hilft einem, wenn man nicht gleich die richtige Antwort weiß.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19SHL02	lobt einen, wenn man etwas richtig macht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19SHL03	geht im Unterricht auf Ideen der Schüler/innen ein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19SHL04	erklärt genau, warum eine Antwort nicht ganz korrekt war.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19SHL05	lässt einem bei Fragen zu wenig Zeit zum Nachdenken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19SHL06	tadelt einen, wenn auch nur eine Kleinigkeit nicht ganz korrekt war.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19SHL07	schimpft Schülerinnen und Schüler vor der ganzen Klasse aus.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>14.</b>	<b>Wie stehst du zum Physikunterricht allgemein? (Pro Zelle bitte nur ein Kreuz.)</b>				
V19MIN01	Ich finde, dass der Physikunterricht meistens Spaß macht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19MKO01	Mit meiner Leistung in Physik bin ich zufrieden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19DRU01	Im Physikunterricht fühle ich mich unter Druck.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19MIN03	Der Physikunterricht ist oft sehr interessant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19MKO02	Im Fach Physik komme ich sehr gut mit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19DRU02	Im Physikunterricht fühle ich mich angespannt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19MIN03	Ich finde den Physikunterricht meistens spannend.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19MKO03	Ich glaube, ich bin im Fach Physik ziemlich gut.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V19DRU03	Ich habe Bedenken, Aufgaben im Physikunterricht gut zu lösen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		Stimme voll zu	Stimme eher zu	Stimme eher nicht zu	Stimme gar nicht zu
		4	3	2	1
15.	Welche Wahlfreiheit hast du im Physikunterricht? (Pro Zeile bitte nur ein Kreuz)				
V18MWW01	Wir können oft mitbestimmen, wie der Physikunterricht abläuft.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18MWW02	Die Lehrperson gibt uns immer genau vor, was wir zu tun haben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V18MWW03	Wir können uns oft die Aufgaben aussuchen, die wir bearbeiten wollen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**16. Zum Schluss noch zwei persönliche Angaben**

S18GESCH	Geschlecht	weiblich	<input type="checkbox"/>	männlich	<input type="checkbox"/>
S18MUTTE	Ist Deutsch deine Muttersprache?	Ja	<input type="checkbox"/>	nein	<input type="checkbox"/>

S18ALTER Alter \_\_\_\_\_

S18NOTEP Welche Note hattest Du im letzten Zeugnis in Physik \_\_\_\_\_

### 8.1.3 Schüler-Fragebogen; Zusätzliche Items der Posterhebung

**16. Wie waren Deine Wahrnehmungen bezüglich des Schülerfeedbacks** (Pro Zeile bitte nur ein Kreuz)

V28WSF01	Die Lehrperson nahm das Schülerfeedback ernst.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28WSF02	Sind dir seit dem Beginn der Feedbacks Veränderungen im Unterricht aufgefallen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28WSF21	Falls du „stimme eher zu“ oder „stimme voll zu“ angekreuzt hast: Welche Punkte sind dir besonders aufgefallen? _____ _____ _____				
V28WSF22	Glaubst du, dass die Veränderungen durch das Feedback ausgelöst wurden?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28WSF03	Ich fand es gut der Lehrperson Feedback geben zu können	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28WSF04	Die Anzahl der Feedbacks in den letzten Wochen war zu hoch.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28WSF05	Ich hätte in den letzten Wochen gern noch häufiger Feedback gegeben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		Vier mal	Drei mal	Zwei mal	Ein mal
		4	3	2	1
17. Wie häufig kam es vor, dass... (Pro Zeile bitte nur ein Kreuz)					
V28FEE01	die Lehrperson das Ergebnis des Schülerfeedbacks der Klasse gezeigt hat?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28FEE02	die Lehrperson mit der Klasse die Ergebnisse des Schülerfeedbacks diskutiert hat?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**18. (V20FRA05) So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich gut, weil....(max. 3 Stichpunkte)**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

**19. (V20FRA06) So ein Schülerfeedback, wie es in den letzten Wochen durchgeführt wurde, finde ich schlecht, weil....(max. 3 Stichpunkte)**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_



## 8.1.4 Lehrpersonen-Fragebogen, Items der Präerhebung

## 1. Wie würden Sie diese Klasse kennzeichnen? Die Schülerinnen und Schüler in dieser Klasse sind...

V1LGL001	interessiert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	uninteressiert
V1LGL002	leistungsfähig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	leistungsschwach
V1LGL003	diszipliniert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	undiszipliniert
V1LGL004	aufmerksam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unaufmerksam
V1LGL005	ausdauernd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	weniger ausdauernd
V1LHOM01	leistungsmäßig homogen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	leistungsmäßig heterogen
V1LGL006	selbständig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unselbständig

## 2. Wie beurteilen Sie die folgenden Aussagen zum Schülerfeedback? (Pro Zeile bitte nur ein Kreuz.)

		stimme voll zu	stimme eher zu	stimme eher nicht zu	stimmt gar nicht zu
		4	3	2	1
V1LESF01	Schülerinnen und Schüler sind, gemittelt auf Klassenebene, kompetente Beurteiler von Unterricht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1LESF02	Das Urteil von Schülerinnen und Schülern bzgl. des Unterrichts muss man ernst nehmen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1LESF03	Ich glaube, dass ein Schülerfeedback zu einer Verbesserung des Unterrichts führen kann.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1LESF04	Schülerinnen und Schüler bewerten vor allem den Unterricht am besten, der für sie am bequemsten ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1LESF05	Es besteht eine große Gefahr, dass Schülerinnen und Schüler das Feedback als Racheinstrument benutzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## 3. In welchem Maße setzen sie folgende Strategien im Unterricht ein? (Pro Zeile bitte nur ein Kreuz.)

		Häufig	Eher Häufig	Eher Selten	Selten
	In meinem Unterricht...	4	3	2	1
V1SAKH01	lasse ich die Schülerinnen und Schüler selbst Wege zur Lösung von Problemen entdecken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH02	lasse ich die Schülerinnen und Schüler ihre eigenen Lösungsideen diskutieren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH03	haben die Schülerinnen und Schüler Gelegenheit in Paaren/Kleingruppen Anwendungsprobleme gemeinsam zu lösen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH04	erfrage ich die Vorstellungen und subjektiven Erklärungen meiner Schülerinnen und Schüler und baue auf diesen auf.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH05	erfasse ich den individuellen Fähigkeitsstand und gebe den SchülerInnen gezielt angepasste Aufgaben und Rückmeldung.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH06	wähle ich Aufgaben so aus, dass sie den Aufbau und die Erweiterung naturwissenschaftlicher Konzepte ermöglichen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH07	unterstütze ich die Schülerinnen und Schüler dabei, ihr eigenes Lernen wertzuschätzen und Vertrauen in ihr Lernvermögen zu entwickeln.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH08	unterstütze ich das forschend-entdeckende Lernen der Schülerinnen und Schüler durch unterschiedliche Methoden und Lernarrangements	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V1SAKH09	lasse ich die Schülerinnen und Schüler neues Wissen durch Experimentieren entwickeln.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

V15K110	gebe ich den Schülerinnen und Schülern die Möglichkeit, aus Fehlern oder falschen Vorstellungen zu lernen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V15K111	erkläre ich neue Inhalte und Zusammenhänge ausführlich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V15K112	gebe ich genaue Anweisungen, wie die Aufgaben zu erledigen sind.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V15K113	achte ich besonders darauf, dass das richtige Fachvokabular benutzt wird.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V15K114	ist es mir wichtig, dass die Schülerinnen und Schüler den durchgenommenen Stoff exakt wiedergeben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V15K115	finden ausführliche Klassengespräche statt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V15K116	ist es mir sehr wichtig mit dem Stoff durchzukommen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**4. Wie würden Sie den Unterricht in dieser Klasse kennzeichnen? (Pro Zeile bitte nur ein Kreuz.)**

		4	3	2	1
		stimme voll zu	stimme eher zu	stimme eher nicht zu	stimme gar nicht zu
V11LGN01	Diese Klasse ist überdurchschnittlich laut.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN02	Diese Klasse ist relativ schwer zu disziplinieren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN03	Ich habe die Klasse gut im Griff.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN04	In dieser Klasse ist es sehr einfach, eine ruhige Lernatmosphäre zu schaffen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN05	Es ist schwer, sich in dieser Klasse durchzusetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN01	Das Lehrer-Schülerverhältnis ist sehr gut.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN02	Ich würde die Klasse vermissen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN04	Ich bin froh, wenn ich die Klasse nicht mehr unterrichten muss.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V11LGN05	Würde ein Kollege/eine Kollegin über die Klasse schimpfen, sähe ich das Problem eher bei dem Kollegen/der Kollegin.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**5. Persönliche Angaben**

LOGGESCH	Geschlecht	weiblich <input type="checkbox"/>	männlich <input type="checkbox"/>
LOGALTER	Alter	29 Jahre oder jünger <input type="checkbox"/> 40-49 Jahre <input type="checkbox"/> 60 Jahre und älter <input type="checkbox"/>	30-39 Jahre <input type="checkbox"/> 50-59 Jahre <input type="checkbox"/>
LOGBERUF	Wie lange sind Sie bereits im Schuldienst als Lehrperson tätig?	1-2 Jahre <input type="checkbox"/> 6-10 Jahre <input type="checkbox"/> 21-30 Jahre <input type="checkbox"/>	3-5 Jahre <input type="checkbox"/> 11-20 Jahre <input type="checkbox"/> mehr als 30 Jahre <input type="checkbox"/>



LOSNUFAE	Welche Fächer unterrichten Sie außerdem? (mehrere Antworten möglich)	<div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px;"></div>	
LOSPENGU	Wie groß ist der Umfang Ihrer Unterrichtstätigkeit insgesamt pro Woche?	1 bis 5 Stunden <input type="checkbox"/> 14 bis 19 Stunden <input type="checkbox"/> 27 und mehr Stunden <input type="checkbox"/>	6 bis 13 Stunden <input type="checkbox"/> 20 bis 26 Stunden <input type="checkbox"/>
LOSPENSK	Wie viele Wochenstunden unterrichten Sie in der untersuchten Klasse?	2 oder weniger <input type="checkbox"/> 3-4 <input type="checkbox"/>	5-6 <input type="checkbox"/> Mehr als 6 <input type="checkbox"/>
LOSVAHNU	Seit wann unterrichten Sie die untersuchte Klasse? Seit...	weniger als einem halben Jahr <input type="checkbox"/> 6-12 Monaten <input type="checkbox"/> 12-18 Monaten <input type="checkbox"/> 18-24 Monaten <input type="checkbox"/>	2-3 Jahren <input type="checkbox"/> 3-4 Jahren <input type="checkbox"/> mehr als 4 Jahren <input type="checkbox"/>
LOSSTUDI	Welche Fächer haben Sie studiert? (mehrere Antworten möglich)	<div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px;"></div>	

## 8.1.5 Lehrpersonen-Fragebogen; zusätzliche Items der Posterhebung

**4. Nun geht es um Ihre Einschätzung, welche Wirkung das Schülerfeedback auf Ihren Einsatz der folgenden Unterrichtsstrategien hatte. Kreuzen Sie bitte an, wie sich die Häufigkeit der Strategien aufgrund des Schülerfeedbacks verändert hat, im Vergleich zum Unterricht vor der Feedbackphase. (Pro Zeile bitte nur ein Kreuz.)**

**In meinem Unterricht...**

		Häufiger in der Feedbackphase	keine Wirkung	seither in der Feedbackphase
		2	1	0
V28KD01	lasse ich die Schülerinnen und Schüler selbst Wege zur Lösung von Problemen entdecken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD02	lasse ich die Schülerinnen und Schüler ihre eigenen Lösungsideen diskutieren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD03	haben die Schülerinnen und Schüler Gelegenheit in Paaren/Kleingruppen Anwendungsprobleme gemeinsam zu lösen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD04	erfrage ich die Vorstellungen und subjektiven Erklärungen meiner Schülerinnen und Schüler und baue auf diesen auf.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD05	erfasse ich den individuellen Fähigkeitsstand und gebe den SchülerInnen gezielt angepasste Aufgaben und Rückmeldung.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD06	wähle ich Aufgaben so aus, dass sie den Aufbau und die Erweiterung naturwissenschaftlicher Konzepte ermöglichen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD07	unterstütze ich die Schülerinnen und Schüler dabei, ihr eigenes Lernen wertzuschätzen und Vertrauen in ihr Lernvermögen zu entwickeln.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD08	unterstütze ich das forschend-entdeckende Lernen der Schülerinnen und Schüler durch unterschiedliche Methoden und Lernarrangements	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD09	lasse ich die Schülerinnen und Schüler neues Wissen durch Experimentieren entwickeln.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD10	gebe ich den Schülerinnen und Schüler die Möglichkeit, aus Fehlern oder falschen Vorstellungen zu lernen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD11	erkläre ich neue Inhalte und Zusammenhänge ausführlich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD12	gebe ich genaue Anweisungen, wie die Aufgaben zu erledigen sind.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD13	achte ich besonders darauf, dass das richtige Fachvokabular benutzt wird.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD14	ist es mir wichtig, dass die Schülerinnen und Schüler den durchgenommenen Stoff exakt wiedergeben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD15	finden ausführliche Klassengespräche statt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V28KD16	ist es mir sehr wichtig mit dem Stoff durchzukommen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## 8.2 Einverständniserklärung der Eltern

Mit meiner Unterschrift gebe ich mein Einverständnis, dass mein Kind an der Untersuchung „Wirkmechanismen von regelmäßigem Schülerfeedback“ teilnehmen darf (nähere Beschreibung der Studie siehe Informationsschreiben).

Mit einer Teilnahme unterstützen Sie ein Projekt, welches Information zur zukünftigen Entwicklung von Unterricht liefern kann.

Falls Sie nicht möchten, dass ihr Kind an der Erhebung teilnimmt, müssen Sie dies nicht eigens zum Ausdruck bringen. Sie müssen in diesem Fall keine Erklärung ausfüllen. Schweigen gilt hier nicht als Zustimmung.

---

(Vor- und Nachname des Kindes)

---

(Ort)

---

(Datum)

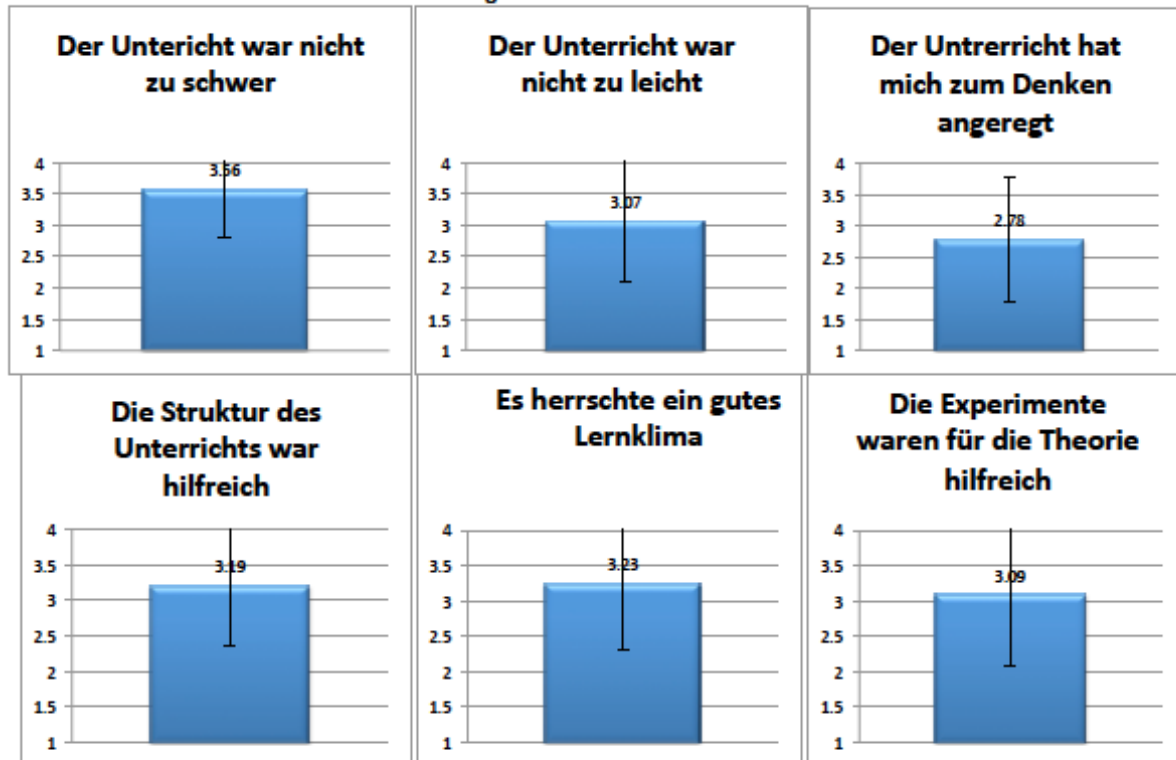
---

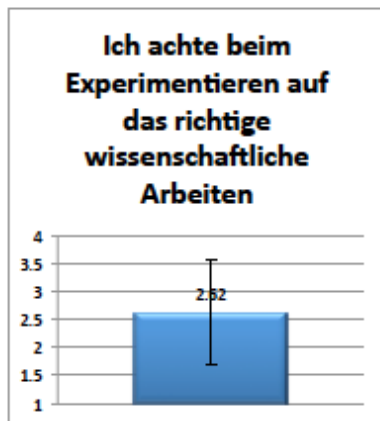
(Unterschrift)

### 8.3 Ergebnisauswertung der Schüler-Kurzfeedbacks für die Lehrpersonen

#### Beispiel für ein ausgewertetes Schüler-Kurzfeedback

Auswertung 4. Schülerfeedback RGG





Folgende Themen sollte man nochmals besprechen

keine

Alle

keine

keine

Anzahl der Schüler/innen,  
die auf die Frage, ob die  
Regeln zum richtigen  
Experimentieren  
eingehalten wurden, nicht  
antworten konnten:

5

---

Was ich noch sagen wollte:

war nice  
Experiment war nice (nais)  
Es war nicht zu laut und wir haben mal einen Versuch gemacht

war richtig Neis

ganz gut  
Der Unterricht war Nais

war gut erklärt

---

## 8.4 Auswertung der Interviews

Auswertungstabelle Interview LP 1			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv		
	emotional negativ		
	nüchtern	2	überraschende Ergebnisse suchen (47, 150); auffällige Ergebnisse suchen (143)
Interpretation	Schüleraussage richtig aber nicht veränderbar	1	Ergebnisse nicht auffällig (zu positiv) (42, 54)
	Schüleraussage richtig und veränderbar	1	Welche Maßnahmen könnten ergriffen werden? (habe ich beim nächsten mal übernommen) (49)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	1	Unterrichtsgeschwindigkeit anpassen (621, 622)
	keine Änderungsabsicht	2	Im Großen und ganzen keine Veränderung vorgenommen (171), Ergebnisse nicht auffällig (42, 54)
Handlung	Handlung verändert	1	Unterrichtsgeschwindigkeit angepasst (621, 622)
	Handlung unverändert		
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht		
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	überraschendes Ergebnis entdeckt (184)
	emotional negativ		
	nüchtern	1	nichtüberraschendes Feedback (153)
Interpretation	Schüleraussage richtig aber nicht veränderbar		
	Schüleraussage richtig und veränderbar		
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	2	Unverstandenes nochmals erklären (421), mehr Zeit für Experimente nehmen (172)

	keine Änderungsabsicht		
Handlung	Handlung verändert	2	Unverstandenes nochmals erklärt (421), mehr Zeit für Experimente (172)
	Handlung unverändert		
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja		
	nein	1	(326)
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert	1	Soziale Kontrolle verbessert (401-405)
	gleich geblieben		
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	(535)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen		
	gleich geblieben		
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen	1	(606)
	gleich geblieben		
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	3	etwas intensiviert, da gemeinsam über Unterricht nachgedacht wurde (524); Individuelle Bedürfnisse werden stärker wahrgenommen (539-545, 577-588), SuS fühlen sich ernster genommen (599)
	gleich geblieben	1	nicht fundamental verbessert (514)
	verschlechtert		
Einschätzung des Verfahrens Schülerkurz-feedback			
Das Verfahren insgesamt	gut	2	es regt zur Kommunikation an (172), gibt Sicherheit (621); Verfahren sinnvoll (677)
	problematisch	2	Feedbackkultur nicht verankert (633, 622), zeitlich knapp (491, 635)
	schlecht	1	Fokus zu wenig auf den Umgang mit den Daten (Diskussion der Daten in der Klasse) (642)
Generelle Einschätzung der Fragebögen	gut	2	Gibt Sicherheit (216); Fragebögen geben guten Anhaltspunkt für die Diskussion (694, 708)
	schlecht	6	Skalen für Diskussion ungeeignet (36, 51, 99); Einzeli-tems sollten sichtbar sein (79); zu häufig (282); Skala kognitive Aktivierung wird von den SuS besser wahrgenommen als es die Fragebogenergebnisse wiedergeben (308), Zeitlich etwas knapp (475); Interpretation der Mittelwerte schwierig (704)
Generelle Einschätzung Vorstellung und Diskus-	gut	2	SuS werden ernst genommen (259); Hier kommen die wichtigen Dinge raus (287)

sion der Ergebnisse in der Klasse	schlecht	1	1. Im 2-Wochen-Rhythmus zeitlich schwierig zu integrieren
In elektronischer Form sinnvoll	ja	2	Mit starkem Fokus auf die Diskussion (676), Werkzeug gut (751)
	nein	1	Technik birgt die Gefahr sinnentleert eingesetzt zu werden (727, 751)
Würde es selbst in elektronischer Form einsetzen	ja		
	nein	1	eher nein, generelle Aversion bzgl. des „sich Verlassens“ auf „Instrumente“ (726-757)
Rhythmus in dem die LP es einsetzen würde	Intervall		
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			LP beklagt häufig die fehlende Feedbackkultur (SuS sind es nicht gewohnt, Vorschläge zur Handlungsänderung mitzugestalten) (u.a. 119, 737). Selbst die Kulturveränderung zu initiieren wäre vor allem zum Ende des Schuljahres hin zu aufwändig (121), Allgemein macht sich die LP viele Gedanken zu den Ergebnissen und wie damit umzugehen sein sollte, Betroffene (SuS) zu Beteiligten machen (u.a. 214-221,649); Das nahe Schuljahresende beschreibt er als hinderlich um mit den Ergebnissen tatsächlich Veränderungen zu initiieren (220) SuS erinnern LP an die Durchführung (465), SuS fanden die durchschnittlichen Ergebnisse zu "der Unterricht hat mich zum denken angeregt" als nicht so schlimm (297, 313)



Auswertungstabelle Interview LP 2			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv		
	emotional negativ		
	nüchtern	1	nicht beeindruckt einfach gemacht (19, 28)
Interpretation	Schüleraussage richtig aber nicht veränderbar		
	Schüleraussage richtig und veränderbar	1	Wo sind schwache Verteilung und bzw. Zustimmung? (52-53)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	1	Mehr Schülerbezogenes Arbeiten bezogen auf Experimentieren (122)
	keine Änderungsabsicht	1	schon lange im Beruf, Änderung schwierig bezogen auf Unterricht allgemein (124)
Handlung	Handlung verändert	1	Schüler mehr selber eigenverantwortlich handeln lassen (133)
	Handlung unverändert		
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	1	Schülerselbstständigkeit erhöhen (Feedback hat daran erinnert) (141, 426-428)
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv		
	emotional negativ	1	Besorgt (256)
	nüchtern	1	Schneller Übergang zur sachlichen Interpretation (217)
Interpretation	Schüleraussage richtig aber nicht veränderbar	1	Ergebnisursache unklar (284)
	Schüleraussage richtig und veränderbar	1	Ergebnis hinterfragt (217)
	Schüleraussage falsch	1	Schüleraussage nicht immer ernst genommen (272)
Motivation	Änderungsabsicht vorhanden	2	Unterricht umstellen und SuS mehr selbständig arbeiten lassen (218-220) Rechnungen wiederholen (315)

	keine Änderungsabsicht	1	Fächerübergreifender Unterricht (266)
Handlung	Handlung verändert	2	Unterricht umgestellt, SuS selbst arbeiten lassen (218-220); Rechnungen wiederholt (315)
	Handlung unverändert	-	-
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja	-	-
	nein	1	LP hat den Unterricht nicht verändert, weil er wusste, dass es ein Feedback stattfinden wird (189)
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert		
	gleich geblieben	1	keine Auswirkung auf Verhalten der SuS (344-351)
	verschlechtert		
	SuS-LP-Verhältnis		
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	(362)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen		
	gleich geblieben	1	genau gleich geblieben (371)
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen		
	gleich geblieben	1	Keine Änderung (385)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert		
	gleich geblieben	1	keine Auswirkungen (357)
	verschlechtert		
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	2	kurzes Verfahren (397), problemlos durchführbar (402)
	problematisch	1	Handlungsvorschläge bei der Rückmeldung fehlen (480-500)
	schlecht		
Generelle Einschätzung der Fragebögen	gut	2	Lenkung der Aufmerksamkeit (151), professionelle Erweiterung der Perspektive (178),
	schlecht	1	SuS können evtl. die eigene Denktivität nicht einschätzen (252)
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	2	Perspektive wird erweitert (162), SuS erfahren Wertschätzung ihres Urteils (166) ,
	schlecht		
In elektronischer Form sinnvoll	ja	1	warum nicht? (441)
	nein	1	evtl. Rücklauf zu gering
Würde es selbst in elektronischer Form einsetzen	ja	1	warum nicht? (441)
	nein		

Rhythmus in dem die LP es einsetzen würde	Intervall		
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			3 x vorgestellt (26) entwicklung der feedbacks mit sus besprochen (27). LP sucht auch ausserhalb des feed-backs das gespräch mit der klasse (152). lp lobt die klasse ausdrücklich (153-157).. LP hat die SuS darauf hingewiesen, dass sie die feedbacks auf die letzten 3-4 Stunden beziehen sollen. (190)

Auswertungstabelle Interview LP 3			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv		
	emotional negativ		
	nüchtern	3	überrascht, dass SuS Unterrichtsqualität besser eingeschätzt haben als die LP (30-36), Ich war nicht aufgeregt, hab mich drauf eingelassen (43) weder enttäuscht noch zufrieden oder unglücklich, es war einfach mal ne Einschätzung, Momentaufnahme, die im Gesamtzusammenhang gesehen werden muss, somit war LP relativ leidenschaftslos (47-52),
Interpretation	Schüleraus-sage richtig aber nicht veränderbar	2	Ergebnisse aus Feedback zu ungenau somit keine Konkreten Anhaltspunkte um Veränderungen durchzuführen (58) SuS sind in einer schwierigen Entwicklungsphase und es lässt sich kaum etwas ändern (209-211)
	Schüleraus-sage richtig und veränderbar	1	Bei Schülerexperimenten ist die Aufmerksamkeit größer und der Unterricht wird besser bewertet (87-94)
	Schüleraus-sage falsch	2	destruktive Kritik (Quatsch) zeigt, dass die Probleme im Unterricht auch bei den SuS zu suchen sind (351-356), Oft sehe ich das anders als die SuS und dann ändere ich nichts (369-372)
Motivation	Änderungs- absicht vorhanden		
	keine Änderungsabsicht	1	Wenn dann nur sehr wenig, merke dass SuS besser arbeiten wenn sie selbst was tun, allerdings sind die Ergebnisse dann fachlich zu ungenau (80-86),
Handlung	Handlung verändert		
	Handlung unverändert	1	nicht gezielt mehr Experimente gemacht (99)
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht		
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv		

	emotional negativ		
	nüchtern	5	Absinken "Kognitive Passung nicht zu schwer" zu MZ 4 aufgrund des schwierigen Themas und schlecht vorbereitet (256-258), Alle anderen Skalen zu MZ 4 blieben auf dem Niveau von MZ 3 weil Aufmerksamkeit der SuS größer war, dies hat sich auf die fast alle Skalen ausgewirkt (274-279), Diejenigen die Interesse zeigen, die waren zufriedener und haben den Schnitt angehoben (285-290), destruktive Kritik aus offenen Kommentaren nehme ich erst mal zur Kenntnis und bin nicht enttäuscht (369, 374) über Verbesserungen nicht wirklich gefreut (419-420)
Interpretation	Schüleraus-sage richtig aber nicht veränderbar	3	Absinken "Kognitive Passung nicht zu schwer" zu MZ 4 aufgrund des schwierigen Themas (256) Alle anderen Skalen zu MZ 4 blieben auf dem Niveau von MZ 3 weil Aufmerksamkeit der SuS größer war, dies hat sich auf die fast alle Skalen ausgewirkt (274-279), Diejenigen die Interesse zeigen, die waren zufriedener und haben den Schnitt angehoben (285-290)
	Schüleraus-sage richtig und veränderbar	2	Den starken Anstieg zu MZ 3 erklärt die LP mit mehr Experimenten und der Zensur des Ergebnisses der SuS durch die LP (68-76, 232) Absinken "Kognitive Passung nicht zu schwer" zu MZ 4 aufgrund schlechter Vorbereitung (258)
	Schüleraus-sage falsch		
Motivation	Änderungs- absicht vor- handen	2	Nach MZ 4 Thema wiederholen mit erweitertem Unterrichtsmaterial (260, 432), Aufgrund von Schüleraussagen zu MZ 1 wurden Themen an- sprechen (327)
	keine Ände- rungsabsicht	1	Nach MZ 1 und 2 war keine Änderungsabsicht vorhanden (209)
Handlung	Handlung verändert	2	Nach MZ 4 Thema wiederholt mit erweitertem Unterrichtsmaterial (260, 432), Aufgrund von Schüleraussagen zu MZ 1 wurden Themen noch- mal angesprochen (327)
	Handlung unverändert	2	Anstieg von "Fokus wissenschaftlichen Arbei- tens" ist nicht auf eine konkrete Handlung der LP zurückzuführen (303), Auf destruktive Kommentare keine Handlung verändert (397)
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler- Feedback folgt?	ja		
	nein	1	(157)
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert	2	manche SuS sind aufmerksamer (482), Disziplin leicht verbessert aber noch nicht gut (491)
	gleich ge- blieben	1	Die die nicht mitgemacht haben, machen das jetzt noch (293)

	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	(449)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen		
	gleich geblieben	1	(459-465)
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen	1	Diejenigen, die es interessier hat haben Alternativvorschläge vorgebracht (470-482)
	gleich geblieben	1	vielen ist das egal (474)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	1	bei ein paar SuS ein bisschen verbessert (440)
	gleich geblieben	1	bei den schweren Fällen gleich geblieben (441)
	verschlechtert		
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	5	Hilfreich, man bekommt Stimmungsbild (504), SuS haben das Gefühl sie werden gehört (506), SuS können Aspekte wie Atmosphäre, Verständlichkeit, eigener Beanspruchung und Hausaufgabenumfang gut beurteilen (518-520), SuS haben die Feedbacks immer ohne Murren ausgefüllt (533-539), Würde nicht viel am Verfahren ändern (543)
	problematisch	1	SuS können nicht beurteilen, ob das Anforderungsniveau hoch genug ist (521-524)
	schlecht		
Generelle Einschätzung der Fragebögen	gut	3	Bei gravierenden Problemen (z. B. Hass) ist ein Fragebogen gut. (136-138), Quatschkommentare zeigen und den SuS den Spiegel vorhalten bzgl. ihrer fehlenden Ernsthaftigkeit (351-362), Es wurden immer die gleichen Fragen gestellt, so konnte man eine Entwicklung sehen (565)
	schlecht	2	Ergebnisse aus Feedback zu ungenau somit keine Konkreten Anhaltspunkte um Veränderungen durchzuführen (58, 544), SuS haben offene Frage (Wdh.) im Fragebogen zu MZ 1 missverstanden: Zeitbezug gesamt statt auf die einzelne Stunde (332),
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	4	Diskussion über die Ergebnisse interessanter (117), Gedächtnisstütze was genau im Unterricht vorgefallen ist (118), SuS wollen wissen was rauskommt (556), Manche SuS sehen, dass eigene negative Extremwerte vom Durchschnitt anders gesehen wird und merken somit, dass der Unterricht evtl. nicht so schlecht war (557-560)
	schlecht		

In elektronischer Form sinnvoll	ja	1	evtl. nicht die ganze Zeit und vor allem mit Klassen, in denen es Schwierigkeiten gibt (579-596)
	nein		
Würde es selbst in elektronischer Form einsetzen	ja	1	evtl. nicht die ganze Zeit und vor allem mit Klassen, in denen es Schwierigkeiten gibt (579-596)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall	2	alle zwei Wochen (601) abhängig von der Stundenzahl evtl auch wöchentlich bei 5 Wochenstunden (605)
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren		3	eher selten intensiver auf Feedback bei der Vorstellung eingegangen. (12) viermal vorgestellt (18). Diskussion beim ersten mal länger, aber hauptsächlich um Missverständnisse des Verfahrens zu klären (13, 22) und weil die LP glaubte das Interesse bei den SuS bzgl. der Ergebnisse war eher gering (25), LP hat dem Feedback eine nicht ganz so große Bedeutung beigemessen, weil sie glaubt auch ohne Rückmeldung erkennen zu können, ob der Unterricht gut gelaufen ist. (48) Den starken Anstieg zu MZ 3 erklärt die LP mit mehr Experimenten und der Zensur des Ergebnisses der SuS durch die LP (68-76), Wenn man weiss, dass es ein Feedback gibt, versucht was "messbares" zu unterrichten (159) Viele soziale Probleme in der Klasse (Mobbing, allgemeine Überforderung, Disziplin, schwierige Entwicklungsphase) (210-225), Klasse bzgl. Arbeitsbereitschaft sehr heterogen (280), LP will sich vermehrt um die kümmern, die bereit sind mitzuarbeiten, SuS sind alt genug (282) LP hat SuS drauf aufmerksam gemacht, dass im offenen Kommentarbereich eine Kritik konkreter sein sollte (398),

Auswertungstabelle Interview LP 4			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv		
	emotional negativ	1	Enttäuscht, dass Alltagsrelevanz des Unterrichts nicht positiv wahrgenommen wurde (38)
	nüchtern	2	Verwundert, da weniger Interesse am Stoff als Gedacht (33), Verwundert über wider erwarten hohe Standardabweichung (48),
Interpretation	Schüleraussage richtig aber nicht veränderbar	2	Mathematik ist nicht beliebt aber nötig (68), mehr Experimente könnten zu Lasten der benötigten Mathematik gehen, Zeitproblem (84,110)
	Schüleraussage richtig und veränderbar	3	zum Teil aber auch veränderbar (67), Experimentieren kommt zu kurz (81), Themenreihenfolge ungünstig (140)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	1	Mehr Schüler-Experimente, die nicht allzu viel Zeit kosten (119)
	keine Änderungsabsicht		
Handlung	Handlung verändert	1	einfache Experimente durchgeführt (123-130)
	Handlung unverändert		
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	3	Themenreihenfolge ändern (134); Unterrichtsteile, die entweder nicht fruchtbar ; waren oder keinen Spaß gemacht haben komplett anders machen oder rausnehmen (147); überlegt sich, mehr Handexperimente durchzuführen, wenn möglich (159)
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	Über Lob gefreut (322)
	emotional negativ		
	nüchtern	2	SCHÜLER-KURZFEEDBACK 3 Einbruch bei "KP nicht zu schwer" aufgrund von mehr Mathe (244), Atmosphärenveränderung in den Stunden wahrgenommen. (295) Änderung kogn. Aktivierung auf Rahmenbedingungen bezogen (266-284)
Interpretation	Schüleraussage richtig aber nicht veränderbar	3	SCHÜLER-KURZFEEDBACK 3 Einbruch bei "KP nicht zu schwer" aufgrund von mehr Mathe: Das ist eben der Lernstoff, der ist nicht verhandelbar (247); Absinken der KA: Von der Schwere des Stoffs abhängig-> leichter Stoff mehr kogn. Aktivierung, Kurz vor Weihnachten viele Klausuren und SuS nicht mehr aufnahmefähig 281-288); Änderung der Unterrichtsatmosphäre: schwer veränderbar, liegt in der Natur der Sache kurz vor Weihnachten (306)



	Schüleraussage richtig und veränderbar		
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	2	Formel werden nochmals wiederholt (345-348), Sinussatz wiederholen (385)
	keine Änderungsabsicht	2	Kognitive Aktivierung: Sequenz lange vorher geplant (315), Lob hatte keine zusätzliche Motivation bewirkt, weil sowieso immer motiviert(335)
Handlung	Handlung verändert		
	Handlung unverändert	1	In Zukunft Formel Wiederholung fest eingeplant (348)
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja		
	nein	1	Ich hab höchstens zeitlich etwas gekürzt (226)
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert		
	gleich geblieben	1	erst verschlechtert dann verbessert, liegt aber an der Gesamtsituation (478-480)
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen	1	Hat den Blick geschärft für Multidimensionalität von Unterricht (409,468)
	gleich geblieben		
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen	2	Einzelne hats zum Denken angeregt (451),
	gleich geblieben	1	Insgesamt keine Veränderung (450)
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen		
	gleich geblieben	1	kann ich nicht Einschätzen (466)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert		
	gleich geblieben	1	SuS-LP- Verhältnis war zuvor schon gut (434-439)
	verschlechtert		
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	2	generell sinnvoll (487), Ankreuzen mit zusätzlichem freiwilligen Kommentar (519)
	problematisch	3	detaillierter wäre besser (491), Einschätzung bei SuS nahm die Motivation das Feedback auszufüllen ab (498) Feedback zu häufig (509)
	schlecht		
Generelle Einschätzung der Fragebögen	gut	2	SuS bleiben anonym (165,441), Jeder muss sich mit dem Unterricht auseinandersetzen (187)
	schlecht		

Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	3	Kommt mehr dabei raus als beim schriftlichen Feedback (183); SuS wollen wissen wie die Ergebnisse sind und auch darauf eingehen (201), Diskussion kann auch dazu führen, dass das Verhalten von Mitschülern kritisch beleuchtet wird (212)
	schlecht	1	Nur ein Teil der SuS beteiligt sich (189)
In elektronischer Form sinnvoll	ja	1	(530)
	nein		
Würde es selbst in elektronischer Form einsetzen	ja	1	Auf jeden Fall, wenn sich wirklich jeder daran beteiligen kann (530)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall		3-4 mal im Jahr nach einer Unterrichtssequenz (515)
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren		1	Eine Schülerin hat auf die störenden Jungs reagiert (302).

Auswertungstabelle Interview LP 5			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv	2	zum Teil erfreut (47), Ergebnisse im durchschnitt so, das ich zufrieden bin (47-51)
	emotional negativ	1	zum Teil enttäuscht (46-47)
	nüchtern	1	versucht es rational zu sehen, was wurde wie gemacht bei welchen Rahmenbedingungen (37-40)
Interpretation	Schüleraussage richtig aber nicht veränderbar	1	SuS können oft die Randbedingungen nicht erkennen (61-63)
	Schüleraussage richtig und veränderbar	1	Es hat eine gewisse Aussagekraft, weil sie auf konkrete Fragen antworten (58-61).
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	3	Routine aufbrechen (80), Stoff langsamer durchnehmen(81) und evtl wiederholen (82)
	keine Änderungsabsicht		
Handlung	Handlung verändert	3	Dinge wurden mehrmals erklärt (84), nachgefragt, ob es wirklich verstanden wurde (85), Aufforderung zum Nachfragen (85-86)
	Handlung unverändert		
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	3	für Ruhe sorgen (91,222), Themen langsamer durchnehmen (92), selbst ein Feedback installieren, das etwa einmal im Halbjahr stattfindet (92-93),
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	gefremt, dass die SuS die Veränderung zur Verbesserung der Atmosphäre bei MZ 2 wahrgenommen haben (220)
	emotional negativ		

	nüchtern	4	Schülerbewertungen hängen stark von dem Stundeninhalt und der Methoden ab (131), Durchschnittliche Werte "kognitive Aktivierung" sind auf hohen Anteil Frontalunterricht zurückzuführen (148-150) Anstieg "Struktur" ist evtl. auf bzw. SuS haben durch das Ausfüllen der Fragebogen den Fokus auf die Struktur gerichtet (193) Anstieg Struktur weil LP das Fragen stellen mehr gefordert hat (197-198), interessante Erkenntnisse, dass den SuS eine legere lockere Art gar nicht so passt (216-217)
Interpretation	Schüleraussage richtig aber nicht veränderbar	1	Schülerbewertungen hängen stark von dem Stundeninhalt und der Methoden ab (131),
	Schüleraussage richtig und veränderbar	1	Elektrizitätsformel wiederholen (235)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	2	Lernatmosphäre verbessern, da es zu laut im Unterricht war (69-71, 207). Hätte auf jeden Fall die Elektrizitätsformel wiederholt (239),
	keine Änderungsabsicht		
Handlung	Handlung verändert	1	Hab mehr Ruhe reingebracht und mehr Ruhe verlangt (74-76),
	Handlung unverändert	1	Aufgrund der Ferien habe ich Elektrizitätsformel nicht wiederholt (236)
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja		
	nein	1	(126)
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert	1	Hat sich verbessert, da LP mehr darauf geachtet hat (207, 220)
	gleich geblieben		
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	(264)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen	1	Während des Ausfüllens der Feedbacks waren die SuS leise und haben ernsthaft die Bögen ausgefüllt, man hat gemerkt, dass es ihnen wichtig ist (269-273)

	gleich geblieben	1	Während des Unterrichts wurden sie nicht als reifer wahrgenommen (274)
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen		
	gleich geblieben	1	da nicht thematisiert wurde (278)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	1	Da ich auf gewisse Wünsche eingehen konnte und gewisse Dinge auch beachtet habe wurde SuS LP Verhältnis verbessert (248-252)
	gleich geblieben		
	verschlechtert		
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	2	Verfahren war in sich stimmig (289), Wenig Zeitaufwand (290)
	problematisch	1	Es fehlt ein Hinweis auf den Fragebögen zu Inhalt und Methoden der Unterrichtsstunden (306-311), manchmal wird Blödsinn bei offenen Fragen reingeschrieben, damit es bei der Diskussion zu sehen ist (96, 296)
	schlecht		
Generelle Einschätzung der Fragebögen	gut	5	offene Fragen haben eine sehr hohe Aussagekraft (94), Feedbackbögen haben mehr zur Unterrichtsentwicklung beigetragen als Vorstellung und Diskussion (107), Skalen wurden gebildet, (285), Anonymität (288), SuS hatten immer Lust die Bögen auszufüllen (294)
	schlecht		
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	1	Die Ergebnisse der Klasse vorstellen halte ich für wichtig (115-116)
	schlecht	2	SuS haben Probleme die Standardabweichung zu verstehen (27-29). Oft war die zu besprechende Unterrichtseinheit zu weit weg (108)
In elektronischer Form sinnvoll	ja	1	würde den Aufwand des Auswertens verringern (321)
	nein		
Würde es selbst in elektronischer Form einsetzen	ja	1	würde den Aufwand des Auswertens verringern (321)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall	1	Bei digitalem Verfahren: möchte es spontan machen, öfters als im momentanen Alltag (333-336) evtl. einmal im Halbjahr bei einem analogen Verfahren (93)

Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			LP hat nicht immer alle Kommentare zu den offenen Fragen vorgestellt (18), LP war motiviert Handlungsmaßnahmen durchzuführen, diese führten nicht zum erhofften Ziel, da SuS lieber raus wollten statt Dinge geklärt zu haben (zeitliche Rahmenbedingungen nicht ausreichend) (86-88),
--	--	--	--

Auswertungstabelle Interview LP 6			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv	1	man nimmt das erst mal immer emotional auf, grundsätzlich war ich zufrieden (35,37,47)
	emotional negativ	3	man nimmt das immer emotional auf, Dinge die nicht so positiv waren, da fragt man sich schon woran das liegen könnte (35,38) ich war enttäuscht es als um die Experimentiergeschichten ging (52-53). Ich habe mich gefragt, warum die keine Ahnung von wissenschaftlichem Arbeiten haben (56-60).
	nüchtern	4	ich war hellhörig als es um die Experimentiergeschichten ging (53), ich bin zufrieden, was die SuS mir zurückgemeldet haben (26) Es gehört zu Lehrerprofessionalität, die Ergebnisse nicht persönlich zu nehmen (41), ich war schon zufrieden, wenn der Balken in der Mitte war (47)
Interpretation	Schüleraussage richtig aber nicht veränderbar	2	Experimentier-Materialien an der Schule sind aus verschiedenen Gründen sehr dürtig vorhanden (77-90,)In bestimmten Bereichen haben sie recht aber Stundenplan, Raum, Situation lässt es nicht zu (71-73)
	Schüleraussage richtig und veränderbar		
	Schüleraussage falsch	1	SuS haben manchmal eine andere Wahrnehmung (68-69)
Motivation	Änderungsabsicht vorhanden	3	Ich bin auf Schülergespräche anders eingegangen (99), Ich lass mehr Schülerantworten so stehen (102-106), SuS sollen merken, dass sich negativ beurteilte Aspekte ändern (177)
	keine Änderungsabsicht		
Handlung	Handlung verändert	2	Ich bin auf Schülergespräche anders eingegangen (99), Ich lass mehr Schülerantworten so stehen (102-106)
	Handlung unverändert		
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	1	Mehr Wert auf das Experimentieren legen (125,131-133)
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	Viel Lob hab ich positiv aufgenommen, motiviert mich weiter gerne in die Klasse zu gehen (365)
	emotional negativ	1	tiefer Werte bei wissenschaftlichem Arbeiten, stört mich, dass der so tief liegt (314)

	nüchtern	5	Skalen kognitive Passung sinken leicht bei der Feedbackphase: könnte evtl. an der Durchführung von Probeaufgaben liegen aber eher natürliche Schwankungen (227-238), mathematischer Anspruch verringerte sich, somit fanden es manche zu leicht (249-257) Schwankung "kognitive Aktivierung": natürliche Schwankung (280-282) Unterdurchschnittliche Werte bei "Kognitive Aktivierung": Fach, das polarisiert, einige wollen mitdenken, einige nicht, zudem ist die Note in Phy. durch die in Chemie und Bio ausgleichbar (288-302) Unterrichts-atmosphäre zu MZ 4 evtl. etwas schlechter, das SuS die, die Klasse verlassen werden etwas Unruhe stiften (342-346)
Interpretation	Schüleraussage richtig aber nicht veränderbar	5	Skalen kognitive Passung sinken leicht bei der Feedbackphase: könnte evtl. an der Durchführung von Probeaufgaben liegen aber eher natürliche Schwankungen (227-238) mathematischer Anspruch verringerte sich, somit fanden es manche zu leicht (249-257), Unterdurchschnittliche Werte bei "Kognitive Aktivierung": Fach, das polarisiert, einige wollen mitdenken, einige nicht, zudem ist die Note in Phy. durch die in Chemie und Bio ausgleichbar (288-302), tiefere Werte bei wissenschaftlichem Arbeiten: Rahmenbedingen (schlechte Materialien, zu große Klassen) (319-320) Unterrichts-atmosphäre zu MZ 4 evtl. etwas schlechter, da SuS die, die Klasse verlassen werden etwas Unruhe stiften (342-346)
	Schüleraussage richtig und veränderbar	1	tiefere Werte bei wissenschaftlichem Arbeiten: es wird zu wenig Wert darauf gelegt (313)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	3	tiefere Werte bei wissenschaftlichem Arbeiten: Da müsste man was machen (318), Will auf Förderung Schülerforderung eingehen (357) Unterdurchschnittliche Ergebnisse "wissenschaftl. Arbeiten": Wollte Versuche durchführen, die das wissenschaftliche arbeiten genau zum Thema machen (179).
	keine Änderungsabsicht		
Handlung	Handlung verändert	2	Unterdurchschnittliche Ergebnisse "wissenschaftl. Arbeiten": Ich habe Versuche durchgeführt und explizit darauf hingewiesen, worauf es beim wissenschaftl. Arbeiten ankommt (179-181), Forderung von SuS: Auf Spannungsverlauf, Wechselstrom Volt Ampere und Watt nochmal eingehen wurde nicht unmittelbar aber kurz vor der Klassenarbeit wiederholt (358)
	Handlung unverändert		



Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja	1	Hat die Unterrichtsgestaltung aufgrund der Tatsache, dass die Stunde Rückgemeldet wird, dahingehend verändert, dass den SuS ein zentraler Erkenntnisgewinn vermittelt wurde, der ansonsten evtl. erst in einer folgenden Stunde zur Sprache gebracht worden wäre (199-212)
	nein		
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert		
	gleich geblieben	1	(450)
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	generell ernst (395-400)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen	1	passen mehr auf z. B. haben sie beim Thema wissenschaftlichen Arbeiten sehr aktiv mitgemacht (422-430)
	gleich geblieben		
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen	1	ein Teil der Schüler (435)
	gleich geblieben	1	bei manchen gleich geblieben (435)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	1	evtl. hat das Feedback-Verfahren noch mehr Vertrauen geschaffen (386-387)
	gleich geblieben		(379-380)
	verschlechtert		
Einschätzung des Verfahrens Schülerkurz-feedback			
Das Verfahren insgesamt	gut	7	Für die SuS ist es motivierend, dass sie mal ihre Meinung sagen dürfen (143-144,) LP hat den Eindruck, dass es für ihn etwas gebracht hat (160). SuS werden ernst genommen (144,155, 162). Kriegen transparent gezeigt, was die Klasse über den Unterricht denkt (145) Interessant für die LP (461) und für SuS hilfreich (462), Mit den ausgewerteten Skalen konnte ich gut umgehen (492)
	problematisch	2	Evtl. könnte eine LP bei schlechten Werten, die Ergebnisse der Klasse aus Angst nicht zeigen (149), aufwändig (463)
	schlecht		
Generelle Einschätzung der Fragebögen	gut	3	Auswertung der Fragebögen hat meine Unterrichtsentwicklung mehr beeinflusst (142). Fragebogen alleine würde etwas bringen aber nur wenn die SuS merken, dass sich etwas ändert (174-175,178), SuS wollten die Feedbacks immer ausfüllen (466-468)
	schlecht	2	Feedbackbogen bezieht sich nur auf eine Stunde (471) insbesondere weil bei einer Einzelstunde, die Zeit um ca. reduziert wird und dann keine Zeit für einen Schülerversuch bleibt (475-481)

Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	2	Eine Diskussion alleine würde auch wirken, vorausgesetzt, es verläuft nicht im Nichts (184-191), Sollte auf jeden beibehalten werden (487-488)
	schlecht		
In elektronischer Form sinnvoll	ja	1	Skeptisch aber befürwortet es (500-513)
	nein		
Würde es selbst in elektronischer Form einsetzen	ja	1	(513)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall		Ferienabschnitten (520)
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			Eindruck: Die emotionalen Reaktionen sind nicht stark ausgeprägt, sowohl positiv als auch negativ, sie hindern die LP nicht daran nüchtern zu interpretieren (35-60, insbesondere 41)

Auswertungstabelle Interview LP 7			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv	1	Ich habe mich über Lob gefreut (273, 281)
	emotional negativ	1	gefrustet nach den dritten Feedback, da keine Veränderung zu erkennen war (80)
	nüchtern	1	Im großen und Ganzen war das okay (35-36, 78)
Interpretation	Schüleraussage richtig aber nicht veränderbar		
	Schüleraussage richtig und veränderbar		
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	1	wollte Stunde anders durchführen, weil ich wusste, dass ein Feedback kommt (182)
	keine Änderungsabsicht	1	Nach dem vierten Feedback habe ich dann den Unterricht wie früher durchgeführt (95-96),
Handlung	Handlung verändert	1	Stunde als eine Einheit betrachtet und durchgeführt, weil LP wusste, dass ein Feedback durchgeführt wurde (182)
	Handlung unverändert	2	Nach dem vierten Feedback habe ich dann den Unterricht wie früher durchgeführt (95-96) In Zukunft wird LP nicht stärker darauf achten, dass die Unterrichtsstunden als geschlossene Einheiten unterrichtet werden (209)
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	2	Ein Feedbackverfahren, das ca. alle vier Wochen stattfindet mit ordentlicher Nachbesprechung (115-121), Mehr über Unterrichtsmethodik nachdenken (135-140)
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv		
	emotional negativ		
	nüchtern	3	Schwächere Skalenwerte bei kognitive Aktivierung, sind damit zu erklären, dass der Unterricht für die allermeisten SuS kognitiv gepasst hat (233-242), Relevanz der Aussage bei Lob mit Vorsicht genießen (274-279), Offene Frage "nicht brüllen" war ich etwas angespannt wegen Weihnachtsfeier aber war wahrscheinlich nicht so schlimm wie es dasteht, SuS möchten auch mal provozieren und übertreiben (303-305, 307-310)
Interpretation	Schüleraussage richtig aber nicht veränderbar	3	Keine Erklärung für Absinken der Skala Struktur zu MZ 3 und 4 (265-269), Schüleraussage offene Frage "Lawine wiederholen" konnte nicht nachvollzogen werden (287-295), Änderung wissenschaftliches Arbeiten hätte ich mehr Fokus drauf gelegt wenn ich mehr Zeit gehabt hätte (254-

			259)
	Schüleraussage richtig und veränderbar	2	Ich habe zweimal versucht aufgrund der Ergebnisse etwas zu verändern (87-93), Experimente stärker anleiten (102-109),
	Schüleraussage falsch	1	Die Schüleraussagen waren nicht so krass wie es auf dem Fragebogen ausgesehen hat, sie provozieren und übertreiben manchmal (307-309)
Motivation	Änderungsabsicht vorhanden	3	Ich habe zweimal versucht aufgrund der Ergebnisse etwas zu verändern (87-93), Experimente stärker anleiten (102-109), Habe versucht wegen "nicht brüllen" die Stunde ruhiger zu händeln (306-307)
	keine Änderungsabsicht	1	"Lawine " wiederholen nicht möglich, da Problem nicht klar erkannt wurde (286-295)
Handlung	Handlung verändert	3	Ich habe zweimal versucht aufgrund der Ergebnisse etwas zu verändern (87-93), Experimente stärker anleiten (102-109) Habe versucht wegen "nicht brüllen" die Stunde ruhiger zu händeln. (306-307),
	Handlung unverändert	1	Aufforderung zu MZ 4 "Magnetismus" wiederholen konnte nicht durchgeführt werden weil die Post-Erhebung in der entsprechenden Stunde stattfand. (341)
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja	1	Hat Unterricht anders entwickelt weil Feedback anstand, mehr Schüleraktivität, kleinschrittiger, Stunde als abgeschlossene Einheit und will nach dem Feedbackverfahren zurück zum Alten (178-185)
	nein		
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert		
	gleich geblieben	1	Da hat sich nicht viel geändert (401)
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen	1	prinzipiell ja (369)
	gleich geblieben		
	gesunken	1	Aufgrund von den konstanten Ergebnisse Zweifel, ob die SuS sich von Stunde zu Stunde Gedanken gemacht haben (371-373)
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen	1	SuS waren motiviert, ohne direkten Benefit, die Feedbacks größtenteils ernsthaft auszufüllen, das habe ich bewundert, für eine 9. Klasse nicht selbstverständlich, das zeigt mir, dass man mit der Klasse was machen kann (381-390)
	gleich geblieben	1	Über das Feedback kommt man jetzt nicht dazu sie ernster zu nehmen (381)
	gesunken		
Wahrgenommene Wert-	gestiegen		

Schätzung der eigenen Arbeit durch die SuS	gleich geblieben	1	(395)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	1	Evtl verbessert weil man sich näher gekommen ist (362-365)
	gleich geblieben		
	verschlechtert		
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	4	würde ein ähnliches Verfahren in Zukunft in größeren Abständen einsetzen (115-121, 128-129), SuS hatten immer Lust das Feedback auszufüllen (412), Das Verfahren hat sich gut ritualisiert in den Unterricht eingefügt (415), Der Verlauf der Grafiken regt zur Änderung an (147-149)
	problematisch	3	Sehr dichte Abfrageintervalle (116,419) SuS haben gegen Schluss das Vorstellen der Ergebnisse dazu genutzt einen drauf zu setzen (Quatschkommentare) (17) Man hat bei den Ergebnissen keine Vergleichswerte (34)
	schlecht		
Generelle Einschätzung der Fragebögen	gut		
	schlecht	1	Fragen sollten mehr gebündelt sein, bzw. lieber alternierend einzelne Skalen abfragen, dass es nicht so viel Zeit kostet (420-424)
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	1	Schüler sehen, dass das was sie tun auch was bewirkt (16,435)
	schlecht		Bei Diskussion würde die Anonymität aufgehoben (65-66), Zeitraubend (67-68)
In elektronischer Form sinnvoll	ja	1	(449)
	nein	1	wenn es aufgezwungen wird (450)
Würde es selbst in elektronischer Form einsetzen	ja	1	(449)
	nein	1	wenn es aufgezwungen wird (450)
Rhythmus in dem die LP es einsetzen würde	Intervall		alle vier Wochen (117)
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			Über Feedbacks wurde nicht diskutiert, nur vorgestellt (30), Versuche des Einwirkens hat wenig Veränderung in den Feedbacks gebracht (40-43),

Auswertungstabelle Interview LP 8			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv	1	Erfreut über die eher positive Rückmeldung (44-59,80)
	emotional negativ		
	nüchtern	3	Grundsätzlich zufrieden mit den Ergebnissen (44) Ergebnisse fand ich interessant, gut und nachvollziehbar (46-48), Bei negativen Bewertungen sehe ich das als sinnvolle Kritik, ich bin nie sauer (228-231)
Interpretation	Schüleraussage richtig aber nicht veränderbar		
	Schüleraussage richtig und veränderbar	1	Offene Antworten auf inhaltliche Aspekte der Stunde (145)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	2	Bestätigt in dem was man tut und bestärkt noch mehr aus sich raus zu gehen und die SuS noch mehr zu motivieren, weil das von den SuS honoriert wird (82-87) Offene Antworten der Feedbacks aufgreifen (145-152)
	keine Änderungsabsicht		
Handlung	Handlung verändert	1	Offene Antworten auf inhaltliche Aspekte der Stunde: Punkte wiederholt (145)
	Handlung unverändert	1	Keine generellen Änderungen aufgrund der Rückmeldungen (140)
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	1	Wieder mehr den SuS die Möglichkeit zur strukturierten und regelmäßigen Rückmeldung zu geben (169,556)
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	Forderung Induktionsspannung wiederholen (412)
	emotional negativ		
	nüchtern	1	Sehr positive Schüleräußerungen zu MZ 4 hat mich überrascht, da ich das selbst in dem Maße nicht wahrgenommen habe (521-525)

Interpretation	Schüleraussage richtig aber nicht veränderbar	3	Anstieg MZ 1 zu MZ2 Theoriebezug Experimente: Zu Beginn nahm das Durchführen der Experimente so viel Raum ein, dass bei den SuS die Kapazität für den Theoriebezug nicht mehr vorhanden war (304-324), Anstieg fast aller Skalen zu MZ 4 (außer wissenschaftl. Arbeiten): Lag an der Natur Stunde: da hoher Alltagsbezug, dafür trat wissenschaftl. Inhalt stärker zurück (358-372), eher etwas tiefere Werte bei kognitive Aktivierung: Es ist etwas viel verlangt, die SuS im Unterricht durchgehende kognitiv zu aktivieren, vor allem bei dem etwas schwierigen Thema E-Lehre (387-398),
	Schüleraussage richtig und veränderbar	3	eher etwas tiefere Werte wissenschaftliches Arbeiten: LP achtet auf wissenschaftliches Arbeiten, macht dies aber nicht zum Fokus der Stunde, läuft nebenher (399-407) Forderung Induktionsspannung wiederholen habe ich erwartet (421). Schüleraussage: Elektromotor wiederholen (447)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden		
	keine Änderungsabsicht	2	Anstieg Atmosphäre ab MZ 3: Klasse war neu und man hat sich mit der Zeit aufeinander eingespielt und die SuS fühlten sich sicher und wohl (297-299), eher etwas tiefere Werte wissenschaftliches Arbeiten: Mein Unterricht hat andere Prioritäten (407)
Handlung	Handlung verändert		
	Handlung unverändert	2	Induktionsspannung sowieso immer wieder wiederholt (426). Elektromotor wiederholen wurde sowieso gemacht (447-452).
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja	1	Das Bewusstsein, dass ein Schüler-Kurzfeedback durchgeführt wird hat die Unterrichtsplanung intensiviert (128-133)
	nein		
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert		
	gleich geblieben	1	(588-594)
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	(553)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen		
	gleich geblieben	1	(563-568)
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen	1	Könnte mir vorstellen, dass die Ursachen für die Qualität der Stunde genauer analysiert wird (573-583)
	gleich geblieben		

	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	1	(537-540)
	gleich geblieben		
	verschlechtert		
Einschätzung des Verfahrens Schülerkurz-feedback			
Das Verfahren insgesamt	gut	2	Das Feedback bringt motivational den SuS viel (223), Gibt den SuS das Gefühl wertgeschätzt und respektiert zu werden (547-550, 679),
	problematisch	4	Nach dem vierten mal war die Besprechung der Kurzfeedbacks etwas langweilig (61), SuS sollten mehr frei antworten können, so erhält man mehr Informationen (65-69), Darstellung schwierig, Variationsbreite sehr schmal und schwer zu interpretieren (94-96, 117-118). Man sollte Vergleichswerte zu den Skalen haben (599-611)
	schlecht	1	Verfahren etwas eindimensional (60)
Generelle Einschätzung der Fragebögen	gut	9	SuS fühlen sich wertgeschätzt, weil sie Feedbacks ausfüllen dürfen (196-200) Sie werden gezwungen über den Unterricht nachzudenken (201-203). Negative Kritik aus den Fragebögen regt mich dazu an Dinge zu verbessern (231-235), der Schüler kann Feedback geben ohne gleich eine Bewertung durch die LP (Mimik, Gestik, Kommentar) zu erhalten (246-247), Anonymität (248-252). Jeder SuS wird gefragt (259), Durch die offene Antwortmöglichkeiten einfacher Ansatzpunkt für Wiederholung zu finden (454-457), Die Skalen, die erhoben wurden sind wichtig für den Physikunterricht (632-634), Skalen Theoriebezug Experimente und wissenschaftliches Arbeiten hilft mir als LP den Fokus darauf nicht zu verlieren (634-639)
	schlecht	2	Offene Antworten manchmal zu unkonkret (457-459) Zu wenig freie Antwortmöglichkeiten (179,632-627, 662)
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	2	Diskussion über Unterricht generell wichtig (254-255), Sollte auf jeden Fall beibehalten werden (684-687)
	schlecht	2	Diskussion eher träge, zu stark auf der pädagogischen Ebene und nicht die Sichtweise der SuS (205-212), fehlende Anonymität (248-256)
In elektronischer Form sinnvoll	ja	1	Das wäre praktisch (707-711)
	nein	1	Daten können insgesamt missbraucht werden (711-725)
Würde es selbst in elektronischer Form einsetzen	ja	1	Aber nur wenn die ich die absolute Kontrolle über die Daten habe (699-710, 734)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall		regelmäßig



Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			Ausfüllen der SCHÜLER-KURZFEEDBACK dauert ca. 15min und geht dann auf ca. 10 min runter (30), , Es machte den Anschein, dass die SuS die letzten Feedbacks nicht mehr so motiviert ausgefüllt haben (647-654)
--	--	--	---

Auswertungstabelle Interview LP 9			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv	1	gute Ergebnisse haben positiv gestärkt (41,422),
	emotional negativ	1	schlechte Ergebnisse: wie kann das sein? (427)
	nüchtern	5	Ergebnisse positiv (19), Ergebnisse der Feedbacks von SuS waren nachvollziehbar (32), gespannt, was die Ursachen für die Ergebnisse sind und wie man sie variieren könnte (32-33), SuS hätten Experimentierregeln kennen müssen, weil ich immer darauf hingewiesen habe (43-44), Kognitive Passung ist gut, wir sind auf dem richtigen Weg (54-58)
Interpretation	Schüleraussage richtig aber nicht veränderbar	2	Die räumlichen Rahmenbedingungen (Räumlichkeit, Ressourcen) schränken ein (74,64), Physikstunde als letzte Stunde ist ungünstig (79,191)
	Schüleraussage richtig und veränderbar		
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	1	Mehr über Unterricht diskutieren (92-93,103)
	keine Änderungsabsicht		
Handlung	Handlung verändert	1	Mehr über Unterricht diskutiert (92-93,103)
	Handlung unverändert	1	Ich habe nichts verändert (98)
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	1	mehr Partnerarbeit und Gruppenarbeit, jedoch bietet Raum nur eingeschränkte Möglichkeiten (116-120)
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	Reaktion auf Lob in der ersten Stunde: natürlich gut wenn man solche Rückmeldung bekommt (318)
	emotional negativ	2	Ich habe gedacht, dass des Hochspannungsexperiment einen höheren Stauneffekt verursacht (192-196), Hätte eigentlich höhere Werte erwartet bei der Struktur (216)
	nüchtern	3	Anstieg "Theoriebezug Experimente" aufgrund des Missverständnisses der SuS, was als Experiment zu werten ist (230-237,264); Anstieg könnte auch durch das spannendste Experiment zu MZ4 verursacht sein, aber eher nicht so wahrscheinlich (244-264), Kommentar zu MZ2 und MZ3 es sei zu laut (343)

Interpretation	Schüleraussage richtig aber nicht veränderbar	1	Kommentar zu MZ2 und MZ3 es sei zu laut: Raumverhältnisse sind schlecht (343-347),
	Schüleraussage richtig und veränderbar	3	SuS können die Frage "wissenschaftliches Arbeiten" nicht beantworten (281), SoS fordert mehr selbständige Arbeit (326), Kommentar zu MZ2 und MZ3 es sei zu laut (343)
	Schüleraussage falsch	1	Anstieg "Theoriebezug Experimente" aufgrund des Missverständnisses der SuS, was als Experiment zu werten ist (230-237, 264)
Motivation	Änderungsabsicht vorhanden	3	SuS können die Frage wissenschaftliches Arbeiten nicht beantworten, soll von LP nochmals angesprochen werden (281), Wollte im Bereich der Rechnungen und Erforschung mehr selbständig arbeiten lassen (326-329), Kommentar zu MZ2 und MZ3 es sei zu laut, Raum-Organisatorische Veränderungen beabsichtigt (343)
	keine Änderungsabsicht	1	Reaktion auf Lob in der ersten Stunde: Hat mich bestärkt, so weiter so zu machen wie bisher (320)
Handlung	Handlung verändert	3	Habe mehr selbständig arbeiten lassen (336), Kommentar zu MZ2 und MZ3 es sei zu laut angesprochen und Sitzordnung verändert (343-348) SuS können die Frage wissenschaftliches Arbeiten nicht beantworten, wird von LP nochmals angesprochen (281-289)
	Handlung unverändert		
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja	1	Änderte kaum den Unterricht, weil er wusste, dass ein Feedback stattfindet, einzelne Versuche wurden jedoch eingeschoben um zu sehen wie sie damit gearbeitet haben (166)
	nein		
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert		
	gleich geblieben	1	(414)
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	(383)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen		
	gleich geblieben	1	(390)
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen		
	gleich geblieben	1	(394, 401)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert		
	gleich geblieben	1	(359-368)

	verschlechtert		
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	2	es hat dazu angeregt sich Gedanken über den Unterricht und die Methodik zu machen (430), SuS mussten sich mit der Thematik Unterrichten beschäftigen (434)
	problematisch		Es fehlen Hilfestellungen als Handlungsanleitung auf entsprechende Ergebnisse (458), Am Schluss letzte Stunde, SuS müssen auf den Bus und füllen es dann nicht so ernsthaft aus (469)
	schlecht		
Generelle Einschätzung der Fragebögen	gut	2	Als Grundlage für die Diskussion (129-136, 156), SuS, die sich nicht gerne öffentlich äußern, bekommen durch die Anonymität eine Stimme (152-154)
	schlecht	1	SuS wurde nicht wirklich deutlich, dass ein Demonstrationsexperiment auch als Experiment zu deuten ist (230-235)
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	3	lässt mehr Kontakt zu SuS zu, was Spaß gemacht hat, was man verbessern kann (139-143), man nimmt sich die Zeit und gibt der Schüleraussage Gewicht und Anerkennung (492) würde es im Verfahren beibehalten (500)
	schlecht		
In elektronischer Form sinnvoll	ja	1	(478-490)
	nein		
Würde es selbst in elektronischer Form einsetzen	ja	1	(518, 529)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall	1	über mehrere Monate jede Stunde (546-548)
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			Zu SCHÜLER-KURZFEEDBACK Messzeitpunkt 2 wurden nur 11 Bögen abgegeben, weil die SuS dringend auf den Bus mussten (197-199), Bleibt bei Fragen bzgl. Änderungsabsichten in allgemeinen Formulierungen, während des Interviews wird deutlich, dass er kaum Änderungsabsichten hat.

Auswertungstabelle Interview LP 10			
Kategorie		Anzahl Nennungen	Beschreibung/Begründung (mit Zeilennummer)
Unterrichtsentwicklung allgemein			
Rezeption	emotional positiv	3	es war wirklich spannend (54-57) und emotional (59) aber es war kaum was böses dabei (64-68)
	emotional negativ		
	nüchtern	1	Sie bewerten so wie sie auch kommunizieren (63-64)
Interpretation	Schüleraussage richtig aber nicht veränderbar	1	Es ist schwer beim Thema Radioaktivität Versuche durchzuführen (80-83)
	Schüleraussage richtig und veränderbar	2	Schwierige Unterrichtsatmosphäre in der Klasse (72-75) SuS wünschen die Wiederholung von mehreren Themen (115-120)
	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	4	Ich habe mehr überlegt, was ich praktisch mehr zeigen kann (85-89) Eine Übungsstunde einfügen (115-120), Mehr Versuche reinbringen (109), Methodenvielfalt ändern (111), Möchte früher eingreifen, um Lernatmosphäre zu verbessern (72-75)
	keine Änderungsabsicht		
Handlung	Handlung verändert	3	Atmosphärenverbesserung durch das Umsetzen von SuS und durch schneller konsequenteres Handeln (108-110) Versuche reinzubringen (110) Methodenvielfalt ändern (111),
	Handlung unverändert		
Motivation für zukünftiges Unterrichtshandeln	Beschreibung der Handlungsabsicht	1	Möchte in Zukunft mehr Feedbacks in der Klasse durchführen (131-133),
Unterrichtsentwicklung (speziell aufgrund von auffälligen Ergebnissen)			
Rezeption	emotional positiv	1	Lob hat ein bisschen die Seele gestreichelt (311)
	emotional negativ		
	nüchtern	1	Hat das Lob sachlich gesehen (305)
Interpretation	Schüleraussage richtig aber nicht veränderbar	1	SuS geben zu MZ 1 und MZ 2 an, dass sie das Thema nicht verstanden haben: Die SuS müssen selbständig (Hausaufgabe) die Formel lernen, dann verstehen sie auch den Inhalt (320-331, 342-345),
	Schüleraussage richtig und veränderbar	1	Verbesserung "Kognitive Passung, nicht zu schwer"(202-211)

	Schüleraussage falsch		
Motivation	Änderungsabsicht vorhanden	3	Verbesserung "Kognitive Passung, nicht zu schwer": zusätzliche Übungsstunden einbauen in denen kleinschrittig gearbeitet werden soll (202-211) Den Verlauf der Stunde mit Uhrzeiten an die Tafel geschrieben. (235-239) Übungsstunde zu Zerfallsreihen einbauen (112-119, 359-363)
	keine Änderungsabsicht		
Handlung	Handlung verändert	3	Verbesserung "Kognitive Passung, nicht zu schwer": Habe zusätzliche Übungsstunden eingebaut in denen kleinschrittig gearbeitet wurde (202-211), Anstieg Struktur: Habe den Verlauf der Stunde mit Uhrzeiten an die Tafel geschrieben (235-239) Zerfallsreihen: eine Wiederholungsstunde durchgeführt, die ich im normalen Alltag nicht umgesetzt hätte (112-119, 359-363)
	Handlung unverändert		
Unterrichtsveränderungen vorgenommen in dem Wissen, dass ein Schüler-Feedback folgt?	ja	1	Hat den Unterricht intensiver vorbereitet (Einbau von Experimenten und mehr Methodenvielfalt) und auf Schülerwünsche stärker eingegangen, weil er wusste, dass es ein Feedback gibt (101-103,183-186)
	nein		
Wahrgenommene Reaktionen der SuS auf das Feedback			
Disziplin der SuS	verbessert	1	(268)
	gleich geblieben	1	(409)
	verschlechtert		
SuS-LP-Verhältnis			
Einschätzung des Beurteilungsvermögens der SuS bzgl. Unterricht	gestiegen		
	gleich geblieben	1	Ich nehme das Urteil der Klasse wirklich ernst (391)
	gesunken		
Allgemeine Einschätzung der SuS bzgl. der Reife	gestiegen		
	gleich geblieben		
	gesunken		
Wahrgenommene Wertschätzung der eigenen Arbeit durch die SuS	gestiegen		
	gleich geblieben	1	(403)
	gesunken		
Allgemeine Einschätzung des SuS-LP-Verhältnisses	verbessert	2	Manchen SuS wurde bewusst, dass sie zu lebhaft sind (378-384), Zu manchen SuS ist die menschliche Beziehungsebene stärker geworden (396)
	gleich geblieben		
	verschlechtert	1	Zwischenmenschliche Atmosphäre litt unter der Verbesserung der Lernatmosphäre (261-276)
Einschätzung des Verfahrens Schülerkurzfeedback			
Das Verfahren insgesamt	gut	1	gutes Instrument (134,137,425)
	problematisch		

	schlecht	1	Diagramme sind für die SuS "zu hoch", die haben direkt abgeschaltet (445-449)
Generelle Einschätzung der Fragebögen	gut	6	anonym (135), hat mehr zur Unterrichtsentwicklung beigetragen als die Diskussion (146), Manche SuS haben es ernsthaft ausgefüllt (148-149) Kommentare waren gut (468-470), Für mich als LP waren die Diagramme und Kommentare interessant (475-477), Konkrete Grundlage für Diskussion (164-166)
	schlecht	1	Manche Schüler fanden es blöd (146-147, 438)
Generelle Einschätzung Vorstellung und Diskussion der Ergebnisse in der Klasse	gut	1	SuS waren bei der Vorstellung des Ergebnisverlaufs der 4 SCHÜLER-KURZFEEDBACK am interessiertesten (35)
	schlecht	3	Die SuS nicht sehr interessiert an den Auswertungsergebnissen (22-26, 463-465), Diskussion bei großen Schulklassen sind eher schwierig (164-168), SuS sind für die Art von Diskussion nicht geschult, sie lernen für Noten (173-178)
In elektronischer Form sinnvoll	ja	1	(484)
	nein		
Würde es selbst in elektronischer Form einsetzen	ja	1	(484)
	nein		
Rhythmus in dem die LP es einsetzen würde	Intervall		1-2 mal im Jahr
Weitere typisierende Aspekte des Verhaltens der Lehrperson auf das Schüler-Kurzfeedbackverfahren			LP hat eine Vorstellung der Feedbacks vergessen (3), Maßnahmen im Unterricht konnten durch den schriftlichen Wunsch der SuS vor anderen SuS besser legitimiert werden (119-126), Skeptisch, was langfristige Änderung in der Unterrichtsentwicklung angeht, weil der Schulalltag viele Zwänge (Klassenarbeit, Zeitnot) mit sich bringt (137-140), ein Schüler gab zu, dass er das Feedback als Racheinstrument benutzt hat (152-154), Ein Schüler gab zu Die LP gut bewertet zu haben, obwohl er nichts verstanden hat (156-158), SuS haben mündlich mitgeteilt, dass sie die Strukturbeschreibung jeder Unterrichtsstunde haben wollen, LP kann und will das jedoch in Zukunft nicht immer machen (238-250),

